1. 有Python程序段如下所示：

import pandas as pd

a = pd.Series({“北京”:36，”上海”:80,”深圳”:27,”广州”:45})

a=a+10

print(a[“上海”])

该程序段输出结果为 （ ）

A．70 B.80 C.9­0 D.上海

1. 有Python程序段如下所示：

import pandas as pd

s = pd.Series([2,3,4])

s=s.reindex([2,1,0])

该程序运行时，s.index[0]中的值是 （ ）

A．2 B.3 C.4 D.0

1. 有Python程序段如下所示：

import pandas as pd

import numpy as np

a=np.array([1,2,3,4]).reshape(2,2)

df=pd.DataFrame(a)

print(df.at[1,1])

该程序段运行后的输出结果为 （ ）

A.4 B.3 C.2 D.1

1. 关于DataFrame的说法，正确的是 （ ）
2. DataFrame对象是一个一维数据结构
3. DataFrame对象一列中的数据类型可以不同
4. DataFrame对象中的每一行是一个Series对象
5. DataFrame对象可修改索引

1.pandas,matplotlib,jieba,numpy都是打包好的，内置了很多功能和函数，让你的python更强大，装备方法是import xxx

2.pandas主要含有Series和DataFrame，后者是前者的升级版，两者都是数据结构

3.熟知dataFrame的.T,sum(),mean(),describe(),head(),tail(),sort\_values()等等

4.分行写和连续的等号是等价的，不要看到一行很长的就慌了

5.理解groupby的意义，通过迭代可视化groupby结构变量

6.和csv，excel联系起来

7.明确类、对象、实例的概念

8.这个知识点在学选考卷至少占一个大题

所谓最头疼的东西——Pandas

5. Excel文件“球员数据”中存储了某篮球运动员服役期间的比赛数据，如图所示，杰尼龟编写了Python程序对该数据进行分析处理

1. 分析改球员各赛季场均得分情况的部分代码如下

import pandas as pd

import matplotlib.pyplot as plt

df = pd.read\_excel(“球员数据.xls”)

df[“场均“]=0.0

for i in range(len(df)):

df.at[I,”场均“]=df.at[I,”得分”]/df.at[i,”出场”]

g1=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ #降序排序，取前10

g1=g1([“赛季”，”场均”])

print(g1)

划线处应填入语句（ ）

1. df.sort\_values(“场均”，ascending=True).top(10)
2. df.sort\_values(“场均”，ascending=False).top(10)
3. df.sort\_values(“场均”，ascending=True).head(10)

df.sort\_values(“场均”，ascending=False).head(10)

6.将DataFrame对象df第三行“品牌”列的值修改为“华为”,则程序代码是 （ ）

A. df.品牌=“华为”

B. df.at[2,”品牌”]=“华为”

C. df.at[3,”品牌“]=”华为”

D.df.at[3,品牌]=“华为”

7.杰尼龟获取了高二全体学生期中考试的技术成绩，数据集格式如图所示。为了分析各班级的成绩情况，它编写了一个Python程序对df对象进行操作，请完善划线处的代码

import pandas as pd

df = pd.read\_csv(“jscj.csv”,encoding=”gbk”) #读取csv文件数据

1. 由于杰尼龟乱喷水，把卡比的信息成绩弄糊了，经核实后发现卡比信息考了36分

result = df [ df.姓名==‘卡比‘].index #获得卡比这行的索引

1. 计算每位同学的技术总分

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. 为了惩罚信息不好好学的同学表演节目，杰尼龟想筛选出信息成绩小于20分的同学，结果保存在一个新的DataFrame对象df1

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. 输出每个班级技术全卷的平均分以及对高二年段所有学生的技术全卷成绩从高到低排序

g=df.groupby(“班级”，as\_index=False)

print(\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)

df2=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

8.有个DataFrame对象score包含了“准考证号”“学校名称”“姓名”“总分”“排名”等数据列，下列语句中，可以以学校为单位，输出各校学生“总分”平均值的是 ( )

A. print(score.groupby(“学校名称”，as\_index=False).mean( ) )

B. print(score.groupby(“总分”，as\_index=False).mean( ) )

C. print(score.groupby(“学校名称”，as\_index=False).排名.mean( ) )

D. print(score.groupby(“学校名称”，as\_index=False).describe( ) )

9.杰尼龟把APP的活跃人数存储在“app.xlsx”文件中，如下图所示，现要将结果进行统计，请回答下列问题：



要把“app.xlsx”第一张工作表的信息导入到s1对象中，并进行统计，实现这个功能的python代码如下，将程序划线处填入合适的代码

import pandas as pd

s1=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ #读取并存储

s1\_sum=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ #计算10月人数之和

s1\_aver=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ #计算11月人数的平均值

g=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ #按应用领域分组统计

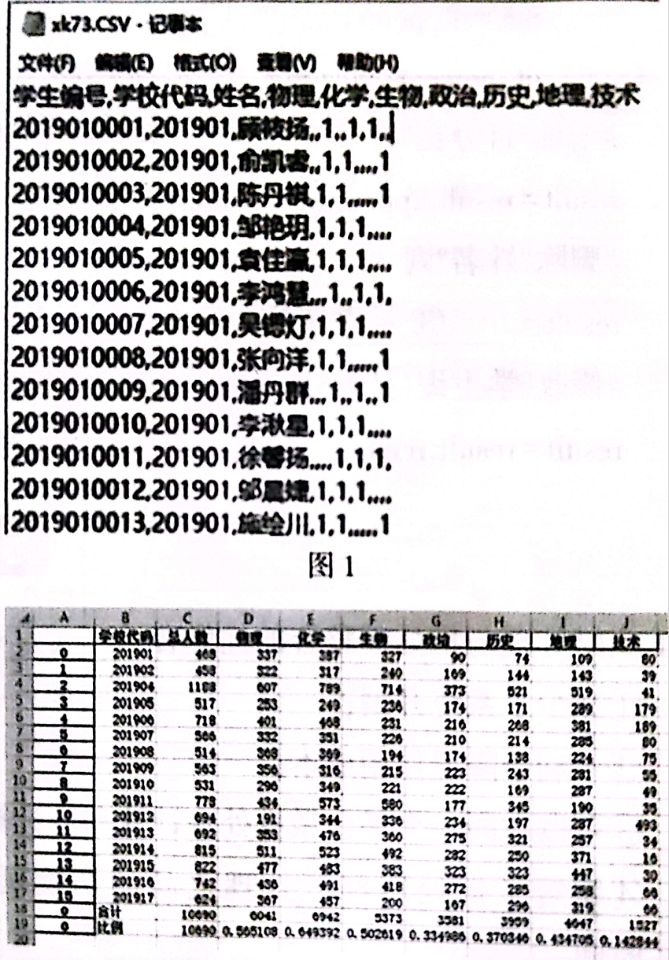
s1\_sort=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ #按11月人数降序排序

print(“12月各领域最多的人数:”,\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_) #输出各领域12月最多的人数

print(\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_)#输出11月人数最多的3条记录

print(\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_) #输出11月人数最少的三条记录

10. 龟龟市选科数据如图1所示，学生从地理、化学、生物等科目中选择三门作为高考选考科目，“1”表示已选择的选考科目。使用Python 编程分析每所学校各科目选考的总人数、全市各科选考总人数及其占比，经过程序处理后，保存结果如图2所示



import pandas as pd

＃读数据到 pandas的DataFrame结构中

df=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_("xk73.csv",sep='.',header='infer',encoding=' utf-8')

km=['物理＇，＇化学＇，＇生物＇，＇政治＇，＇历史＇，＇地理＇，＇技术＇］

zrs=len(df.index)

＃按学校分组计数

sc=df.groupby('\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_’,as\_index=False).count( )

＃对分组计数结果进行合计，合计结果转换为DF结构并转置为行df\_sum=pd.DataFrame(data=sc.sum()).T

df\_sum[‘学校代码’］＝＇合计’ ＃增加＂合计＂行

result=sc.append(df\_sum)

＃百分比计算

df\_percent=df\_sum

df\_percent[‘学校代码’］＝’比例’

for k in km:

per=df\_percent.at[0,k]/zrs

df\_percent[k]=per

＃增加＂百分比＂行

result=result.appeend(df\_percent)

＃删除＂姓名”列

result=\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

＃修改＂学生编号＂为＂总人数＂

result=result.rename(columns={'学生编号’：’总人数’})

#保存结果，创建Excel文件

Result.to\_eccel(“学校人数统计.xlsx”)

加框处语句的作用是\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1，下列属于人工智能的是： （ ）

A．扫地机器人 B.无人驾驶汽车 C.人脸识别 D.网上游戏

2. 下列不属于人工智能领域的是 （ ）

A．机器证明 B.无人驾驶汽车 C.人机博弈 D.伪代码编写

3. 下列不属于人工智能的是 （ ）

A.机器学习 B.模式识别 C.自动控制 D.无人驾驶

4. 下列不属于人工智能主义的是 （ ）

A．符号主义 B. 资本主义 C.联结主义 D.行为主义

5. 1959年华裔科学家王浩用计算机证明了《数学原理》中全部150条一阶逻辑以及200条命题逻辑定理，其应用的人工智能属于 （ ）

A．符号主义 B.联结主义 C.行为主义 D.仿真主义

6. 通过三段论应用的人工智能方法是 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

7. 深度学习可以基于多层神经元的自编码神经网络，可以逐层提取特征，这种人工智

能属于 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

8. 下列描述体现了大数据应用的是： （ ）

A. 杰尼龟货比三家，使用淘宝、京东等平台搜素相同产品，比较价格

B. 电商平台会根据用户的浏览信息进行分析，推送符合爱好的产品和信息

C. 烟雾报警器根据烟雾浓度发出警报

D. 网上订阅的信息会定时地发送到电子邮箱

9. 下列数据服务中运用了医疗大数据的是 （ ）

A. 无人商店 B.股价预测 C.智能投资顾问 D.临床决策支持

1. 判断AI，原来人在做，现在机器能做
2. 符号主义，联结主义，行为主义
3. 人工智能发展历程
4. 深度学习，神经网络，自然语言处理
5. Volume数量,Velocity速度,Variety多样,Value价值

酷但是简单——AI和BIGDATA

10. 下列运用到交通大数据的是\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

①远程医疗 ②超速行驶预警 ③绿化自动灌溉 ④高速道路费用自动收费 ⑤道路口交通情况的视频监控 ⑥车道偏离预警

11.下列不属于大数据的是 （ ）

A. 某超市根据需求和库存情况，对700万商品实时调价

B. 微软大数据成功预测第86届奥斯卡金像奖24个奖项中的21个

C. 中医智慧养生系统根据用户体验情况给出建议和初期诊断

D. 教师分析某位学生历次考试成绩，指出存在的问题和提出改进建议