**《数据挖掘》**

**实验报告1**

**班 级 信计170**

**学 号 10172971**

**姓 名 周嵘鑫**

**批阅教师 钱夕元**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**评分标准：**

1. 成绩分为A/B/C/D四个等级（A为最高）
2. A：实验结果正确；过程进行了清晰的描述；作图清晰美观；对结果进行了详细的分析、讨论
3. B：实验结果基本正确；过程进行了部分描述；作图基本正确；对结果进行了简单的分析、讨论
4. C：实验结果部分正确；过程描述很少；作图不够正确；对结果的分析不够正确
5. D：实验结果出现明显的错误；过程混乱；没有结果呈现；对结果没有分析或分析完全错误
6. 没有按时提交作业、补交作业，成绩降一档
7. 抄袭作业成绩为零分。

**教师评语：**

|  |
| --- |
|  |

**实验成绩：**

|  |  |
| --- | --- |
| 实验名称：Python运行平台的安装与基本操作 | 实验地点：八教611 |
| 所使用的工具软件及环境：Anaconda | |
| 1. 实验目的：    1. 熟悉Python运行环境，安装anaconda，选定Python IDE工具    2. 熟悉Python的基本操作，包括数据结构、数据文件操作    3. 完成实践案例：白葡萄酒品质探索； | |
| 1. 实验内容   （要求参照Python基础（上机操作）PPT内容及实践练习和实践案例要求，完成下述实验操作）   * 1. Anaconda安装   2. Python基本操作   变量、注释、print函数、数据类型、算术运算符、类型转换   * 1. Python数据容器   列表、元组、字典、集合  完成相应的实践练习   * 1. 完成实践案例：白葡萄酒品质探索 | |
| 1. 操作步骤   b)  kobe\_list = [2, 'Jump\_shot', 'Los Angeles Lakers']  print(kobe\_list,kobe\_list.append('POR'))  print(len(kobe\_list))  three\_five = kobe\_list[3:5]  print(three\_five)  kobe\_list = ['Los Angeles Lakers', 'LAL @ POR', 'Jump Shot', 'POR', 'Left Side(L)', '2000-10-31']  # print(kobe\_list)  three\_five = kobe\_list[3:6]  print(three\_five)  year = kobe\_list[-1][0:4]  print(year)  kobe\_tuple = ([2, 'Jump Shot'], 'Los Angeles Lakers', 'POR')  kobe\_tuple[0][0] = 3  print(kobe\_tuple)  kobe\_tuple[0].append(10)  print(kobe\_tuple)  shot\_id = [1,2,3]  shot\_zone\_area = ['Right Side(R)', 'Left Side(L)', 'Left Side Center(LC)']  kobe\_dict = {}  for key, value in zip(shot\_id,shot\_zone\_area):  kobe\_dict[key] = value  print(kobe\_dict)  # print(kobe\_dict.has\_key(5))  print(kobe\_dict.keys())  print(kobe\_dict.values())  print(set(['D', 'R', 'D','E']))  s1 = set(['D','T','R','D','E'])  s2 = set(['D','R','U','W'])  print(s1|s2)  print(s1&s2)  print(s1-s2)  print(s1^s2)  kobe\_tuple = ([2, 'Jump shot'], 'Los Angeles Lakers', 'POR')  kobe\_tuple[0][0] = 3  kobe\_tuple[0][1] = 'Slam Dunk Shot'  print(kobe\_tuple)  shot\_id = [1,2,3]  shot\_zone\_area = ['Right Side(R)', 'Left Side(L)', 'Left Side Center(LC)']  kobe\_dict = {}  for key, value in zip(shot\_id,shot\_zone\_area):  kobe\_dict[key] = value  print(kobe\_dict)  print('1',kobe\_dict.keys())  kobe\_dict.pop(2)  print(kobe\_dict)  kobe\_dict.clear()  a = [2,5,7,4,5,2,10,67,84]  b = []  for x in a:  b.append(x\*2)  print(b)  l = [i\*\*2 for i in range(3,6)]  print(l)  complete = [2,5,4,6,4,3,2,6,5,6]  mean = sum(complete)/len(complete)  delta = [i - mean for i in complete]  print(delta)  complete = ['2','5','7','3','5','2','5','8']  d = {ch:i for i,ch in enumerate(complete)} //生成的字典的value为这个key最后出现的位置的下标  print(d)  f = open('test.txt',mode='w')  # print(f.read())  f.write('this is written in\n')  f.write('7')  f.close()  with open('test.txt') as fi:  content = fi.read()  print(content)  path = '../test\_csv.txt'  f = open(path,mode='r')  con = f.readlines()  print(con[0])  print(con[1])  con\_new = []  for c in con:  temp = c.strip()  temp = temp.split('\t')  con\_new.append(temp)  print(con\_new[0])  print(con\_new[1])  f = open(path,mode='w')  content = []  for x in con\_new:  content.append(x)  # content.append(['1'])  print(content)  # f.close()  for x in content:  x = '\t'.join(x)  x = x + '\n'  f.write(x)  f.close()  白葡萄酒实验：  1、导入数据      2、查看品质类别个数      3、划分数据集      4、统计每个品质的样本量      import csv  path = 'white\_wine.csv'  f = open(path,'r')  reader = csv.reader(f)  content = []  for row in reader:  content.append(row)  f.close()  # for i in range(5):  # print(content[i])  qualities = []  for row in content[1]:  qualities.append(int(row[-1]))  unity\_qualities = set(qualities)  print(unity\_qualities)  content\_dict = {}  for row in content[1:]:  quality = int(row[-1])  if quality not in content\_dict.keys():  content\_dict[quality] = [row]  else:  content\_dict[quality].append(row)  print(content\_dict.keys())  mean\_tuple = []  for key, value in content\_dict.items():  sum\_ = 0  for row in value:  sum\_ += float(row[0])  mean\_tuple.append((key, sum\_/len(value)))  for x in mean\_tuple:  print(x) | |