**实 验 报 告**

**学 院：数学与计算机科学学院**

**课程名称：统计分析**

**专业班级： 智能科学与技术**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 实验项目 | 实验一 | 实验二 | 实验三 | 实验四 | 实验五 | 实验六 | 实验七 | 实验八 | 实验九 | 实验十 | **总评**  **成绩** |
| 评分 |  |  |  |  |  |  |  |  |  |  |  |

**姓 名：**

**学 号：**

**学生实验报告**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 实验项目 | 熟悉python软件实现频数分布的表格操作 | | | |
| ▇必修 □选修 | | □演示性实验▇验证性实验 □操作性实验 ▇综合性实验 | | |
| 实验地点 | 10-315 | | 实验仪器台号 |  |
| 指导教师 | 刘洪久 | | 实验日期及节次 |  |

**一、实验目的及要求：**

1、**目的**

熟悉python、Jupyter软件并掌握其应用，利用其对数据进行次数分布。

2、**内容及要求**

（1）熟悉python数据操作的基本技能。

（2）根据《统计学》教材p38页上面50名工人的原始数据，编写python程序实现表3-7、3-8、3-9、3-10、3-11，保存为.xlsl或.csv文件。

（3）代码不得截屏，保存后的.xlsl或.csv文件结果需截屏展示。

注意：.xlsl或.csv必须通过python建立、编辑、保存，严禁采用excel向.xlsl或.csv输入数据。

**二、仪器用具：**

|  |  |  |  |
| --- | --- | --- | --- |
| **仪器名称** | **规格/型号** | **数量** | **备注** |
| 计算机 |  | 1 |  |
| Python 软件 |  | 1 |  |
| Jupyter软件 |  | 1 |  |

**三、实验资料**

《Python统计分析》

**四、实验注意事项**

1．提前复习和实验内容相关的理论。

2．准确输入实验数据。

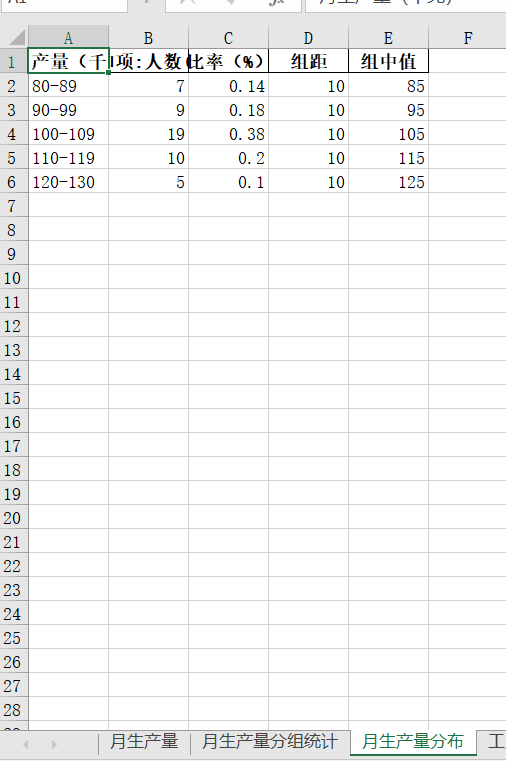
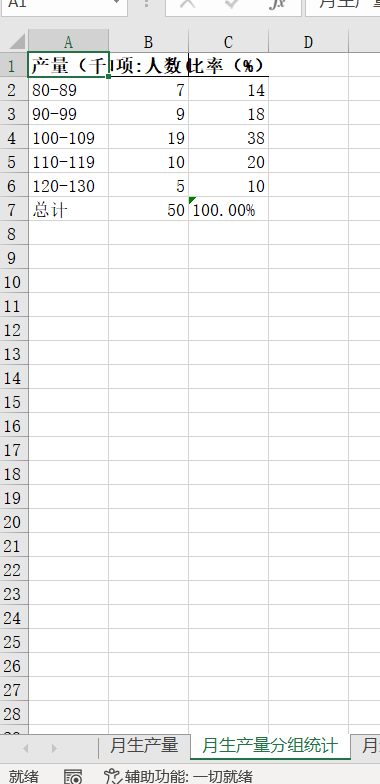
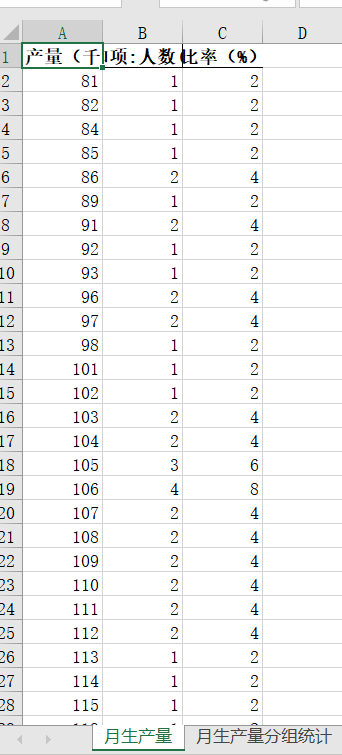
3．熟悉代码语句，准确输入代码。

**五、实验步骤（代码）**

import pandas as pd  
import openpyxl  
  
*#解决数据输出时列名不对齐的问题*pd.set\_option('display.unicode.east\_asian\_width', True)  
  
data = [  
 106, 81, 98, 111, 91, 107, 86, 105, 93, 106,  
 82, 108, 114, 122, 109, 104, 125, 103, 113, 102,  
 106, 84, 128, 104, 91, 112, 85, 96, 115, 89,  
 97, 105, 92, 111, 107, 97, 105, 124, 106, 86,  
 96, 110, 112, 103, 108, 110, 109, 125, 101, 119  
]  
  
*# 统计数量*count\_data = pd.Series(data).value\_counts().sort\_index()  
  
*# 第一个表格*total = len(data)  
percentage = (count\_data / total \* 100).round(2)  
  
table1 = pd.DataFrame({  
 '月生产量（千克）': count\_data.index,  
 '求和项:人数(人)': count\_data.values,  
 '比率（%）': percentage.values  
})  
  
table1.loc['总计'] = ['总计', total, '100.00%']  
  
print("第一个表格: 月生产量（千克）")  
print(table1)  
  
*# 第二个表格 - 分组统计*bins = [80, 90, 100, 110, 120, 130]  
labels = ["80-89", "90-99", "100-109", "110-119", "120-130"]  
grouped\_data = pd.cut(data, bins=bins, labels=labels, right=False)  
group\_counts = grouped\_data.value\_counts().sort\_index()  
  
group\_percentage = (group\_counts / total \* 100).round(2)  
  
table2 = pd.DataFrame({  
 '月生产量（千克）': group\_counts.index,  
 '求和项:人数(人)': group\_counts.values,  
 '比率（%）': group\_percentage.values  
})  
  
table2.loc['总计'] = ['总计', total, '100.00%']  
  
print("\n第二个表格: 月生产量分组统计")  
print(table2)  
  
*# 第三个表格 - 分布统计*group\_midpoints = [85, 95, 105, 115, 125]  
  
table3 = pd.DataFrame({  
 '月生产量（千克）': group\_counts.index,  
 '求和项:人数(人)': group\_counts.values,  
 '比率（%）': group\_percentage.values / 100,  
 '组距': [10] \* len(group\_counts),  
 '组中值': group\_midpoints  
})  
  
print("\n第三个表格: 月生产量分布")  
print(table3)  
  
*# 第四个表格 - 累积次数表*cumulative\_up = group\_counts.cumsum()  
cumulative\_down = total - cumulative\_up + group\_counts  
  
table4 = pd.DataFrame({  
 '月生产量（千克）': group\_counts.index,  
 '人数（人）': group\_counts.values,  
 '比率（%）': group\_percentage.values,  
 '向上累积人数（人）': cumulative\_up.values,  
 '向上累积比率': (cumulative\_up / total).round(2),  
 '向下累积人数（人）': cumulative\_down.values,  
 '向下累积比率': (cumulative\_down / total).round(2)  
})  
  
print("\n第四个表格: 工人月生产量累积次数表")  
print(table4)  
  
with pd.ExcelWriter('月生产量.xlsx','openpyxl') as f:  
 table1.to\_excel(excel\_writer=f,sheet\_name='月生产量',index=False)  
 table2.to\_excel(excel\_writer=f,sheet\_name='月生产量分组统计',index=False)  
 table3.to\_excel(f, sheet\_name='月生产量分布', index=False)  
 table4.to\_excel(f, sheet\_name='工人月生产量累积次数表', index=False)

**六、实验结果及分析**

**Excel表格**

**图形用户界面, 应用程序, 表格, Excel

描述已自动生成**

**数据表格**

第一个表格: 月生产量（千克）

月生产量（千克） 求和项:人数(人) 比率（%）

0 81 1 2.0

1 82 1 2.0

2 84 1 2.0

3 85 1 2.0

4 86 2 4.0

5 89 1 2.0

6 91 2 4.0

7 92 1 2.0

8 93 1 2.0

9 96 2 4.0

10 97 2 4.0

11 98 1 2.0

12 101 1 2.0

13 102 1 2.0

14 103 2 4.0

15 104 2 4.0

16 105 3 6.0

17 106 4 8.0

18 107 2 4.0

19 108 2 4.0

20 109 2 4.0

21 110 2 4.0

22 111 2 4.0

23 112 2 4.0

24 113 1 2.0

25 114 1 2.0

26 115 1 2.0

27 119 1 2.0

28 122 1 2.0

29 124 1 2.0

30 125 2 4.0

31 128 1 2.0

总计 总计 50 100.00%

第二个表格: 月生产量分组统计

月生产量（千克） 求和项:人数(人) 比率（%）

0 80-89 7 14.0

1 90-99 9 18.0

2 100-109 19 38.0

3 110-119 10 20.0

4 120-130 5 10.0

总计 总计 50 100.00%

第三个表格: 月生产量分布

月生产量（千克） 求和项:人数(人) 比率（%） 组距 组中值

0 80-89 7 0.14 10 85

1 90-99 9 0.18 10 95

2 100-109 19 0.38 10 105

3 110-119 10 0.20 10 115

4 120-130 5 0.10 10 125

第四个表格: 工人月生产量累积次数表

月生产量（千克） 人数（人） ... 向下累积人数（人） 向下累积比率

80-89 80-89 7 ... 50 1.00

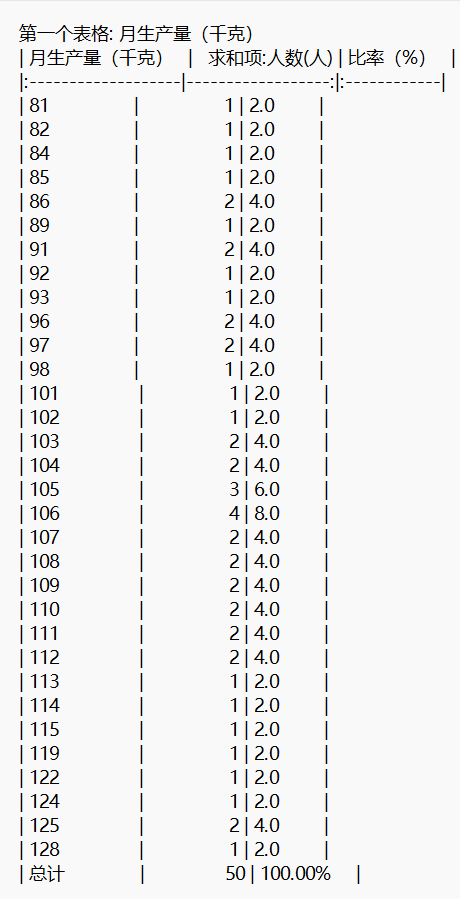
90-99 90-99 9 ... 43 0.86

100-109 100-109 19 ... 34 0.68

110-119 110-119 10 ... 15 0.30

120-130 120-130 5 ... 5 0.10

[5 rows x 7 columns]

****

**图片包含 日历

描述已自动生成**