1. 读取《百合网速配数据》.csv数据和生活标准.csv数据
2. 对《百合网数据》对象最低薪水填充生活标准数据中对象家长的最低生活标准数据，没有的地区或省市以最低生活标准数据的平均值填充。
3. 对《百合网数据》对象最高薪水全是0的数据填充当前列的众数。
4. 对《百合网数据》对象平均薪水全是0的数据填充当前列的平均值。
5. 删除〈百合网数据〉年龄不满20岁的数据对应的数据项，并指明数据项的条数。
6. 对百合网数据的年龄段进行划分，20-30,30-40,40-50,50岁以上的人士对对象的年龄要求，在哪一个年龄段，年龄段的划分仍以20-30,30-40,40-50,50以上为准。
7. 对百合网数据的年龄段进行划分，20-30,30-40,40-50,50岁以上的人士对对象的家乡要求是否一定和自己是老乡，有多少数据是吻合自己的家长，画柱状图显示。
8. 对百合网数据的年龄段进行划分，20-30,30-40,40-50,50岁以上的人士对对象的住房、最低身高、最高身高、平均身高、最高薪水、最低薪水、学历有什么样的要求，分别统计出数据。
9. 分析找对象人士使用自我介绍时使用最多的前20个词汇。
10. 分析数据中离异的找对象人士对对象婚姻程度要求的比率。画出饼图。v
11. 如果现在有一个人，对对象的期望特征如下。

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 对象年龄 | 对象身高 | 对象学历 | 对象薪水 | 对象家乡 | 对象婚姻 | 对象住房 | 最低工资 | 最高工资 | 平均工资 | 对象最低年龄 | 对象最高年龄 | 对象最低身高 | 对象最高身高 |
| 25 | 1.65 | 不限 | 6000 | 上海 | 未婚 | 无 | 1800 | 6000 | 4800 | 23 | 30 | 160 | 170 |

其自己的特征如下：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 年龄 | 身高 | 学历 | 工资 | 家乡 | 婚姻 | 住房 | 自我介绍 |
| 26 | 1.65 | 高中 | 5000 | 湖南 | 未婚 | 无 | [ '希望', '找个', '很帅', '爱'] |

对于这样的特征，目前的数据，其是否可以在当前平台内找到对象，预测一下。

1. 对使用目前的数据进行找对象的预测，其准确率是多少?说一下用的是什么样的算法。
2. 对被统计的找对象人群按薪资划分群体，5000以下，5000-10000，10000-2000，20000以上，用柱形图画出其分布。