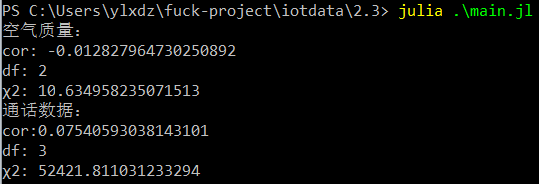
作业2.3 – 1133730117 - 张实唯

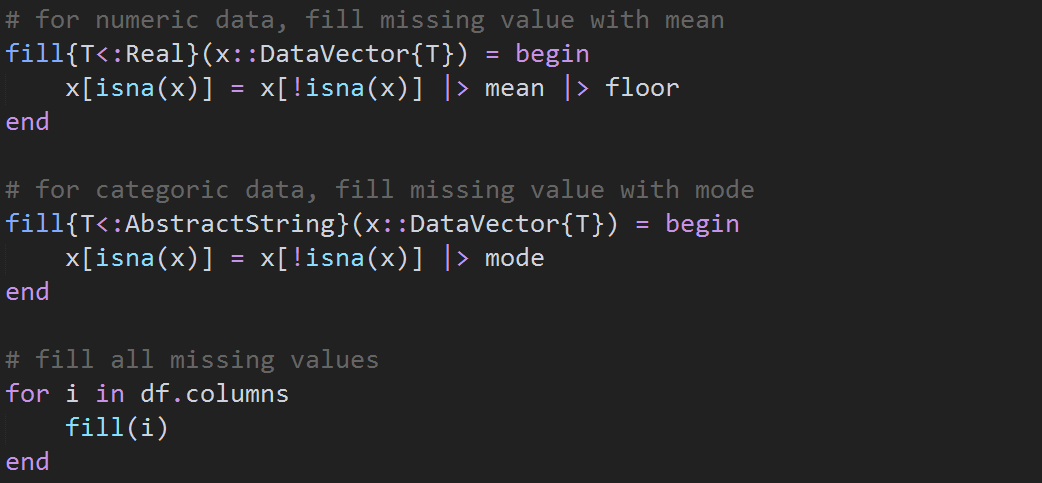
## 实验结果



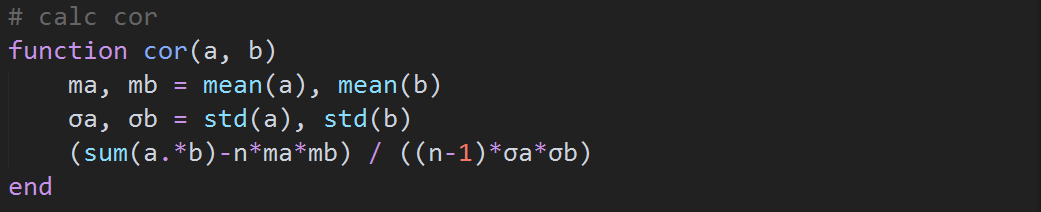
1. 空气质量中的两个变量相关系数的绝对值小于0.05，可以认为二者之间不存在线性相关关系。
2. 空气质量中的两个变量的χ2值为10.63，查表得知自由度为2的χ2分布  
    P(χ2>10.597)=0.005  
   因此拒绝不相关假设，两个变量存在相关关系。
3. 通话数据两个变量相关系数绝对值小于0.1，可以认为二者之间不存在线性相关关系。
4. 通话数据中的两个变量的χ2值为52421.8，经查表得知自由度为3的χ2分布  
    P(χ2>12.8)=0.005  
   因此拒绝不相关假设，两个变量存在相关关系。

## 代码解释

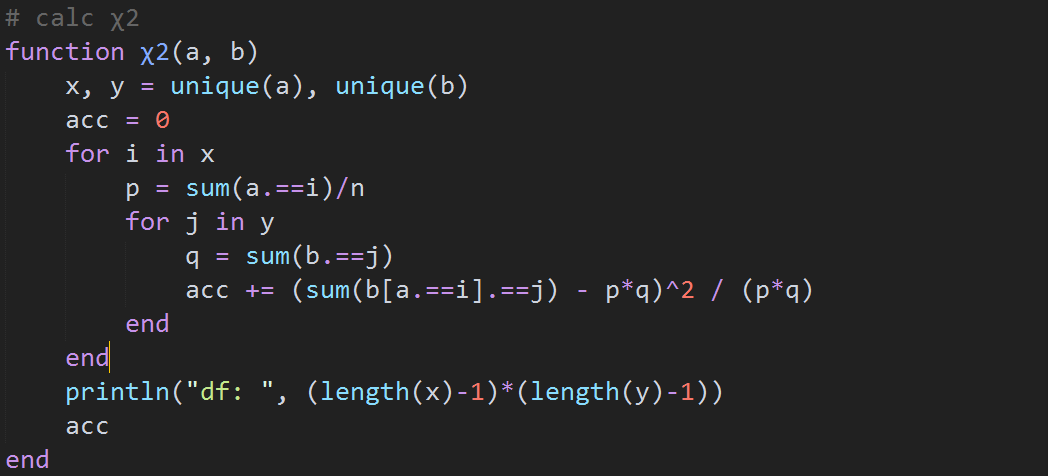
1. 缺失值填充方法: 数值类型使用平均数，标签类型使用众数。



1. 相关系数计算方法: 计算平均数和标准差，然后套公式。



1. χ2计算方法: 统计出现的值计算自由度，然后遍历所有组合来计算χ2。(julia语言支持使用Unicode字符定义函数名)



1. 对数据的分类处理: 计算χ2之前先对数值型数据进行分类。

