**工作任务三数据清洗**

1、我们从各种渠道获得的数据会存在数据不规范、错误等问题，以下哪种情况属于数据值问题？（ ）

A. 日期格式不一致

B. 名字顺序不一致

**C. 数据值缺失**

D. 字段命名不清

2、在问题数据中，以下哪种情况不属于格式问题？（ ）

A. 拼写不一致

B. 名字顺序不一致

**C. 缺失数值被零补位**

D. 日期格式不一致

3、在数据量很大的情况下，哪种数据清洗方式不是最优选择？（ ）

A. 使用Excel表格进行数值处理和函数计算

B. 使用Open Refine软件进行数据清洗

C. 使用计算机语言相关软件，如Python或R软件

**D. 手动编辑**

4、以下哪些软件不能用于数据清洗？（ ）

A. Open Refine软件

**B. PS软件**

C. Excel软件

D. R软件

5、在Open Refine中，进行数据清洗的第一步通常是什么？（ ）

**A. 导入数据文件**

B. 手动编辑每条数据

C. 使用正则表达式替换数据

D. 导出清洗后的数据