文本

描述已自动生成

# 数据治理设计说明

姓名：李琰朕

学号：201830661250

班级：软件三班

年级：2018级

指导教师：汤峰

报告时间：2021年7月3日

## 物理模型设计：

图示

描述已自动生成

如图所示，在登录数据湖的数据治理板块后，选择物理模型，即可看到以创建的“电商模型”，打开后就如上图所示，卖家和商品是n-1的关系，商品和买家之间也是n-1的关系。

## 数据表的字段设计：

### 用户表：

电脑萤幕的截图

描述已自动生成

其中User-id为主键

### 商品表：

电脑截图

中度可信度描述已自动生成

### 卖家表：

图形用户界面, 文本, 应用程序

描述已自动生成

## 需求设计：

### 需求一：

针对按月查看总体销售额和新增注册用户，可以采用如下sql语句：

手机屏幕的截图

描述已自动生成

查询后的结果如图所示：

图形用户界面, 应用程序

描述已自动生成手机屏幕的截图

描述已自动生成

### 需求二：

按月查看城市的销售额：可以使用以下sql语句进行查询并获得以下查询结果：

文本

描述已自动生成

屏幕上有字

描述已自动生成

### 需求三：

按月查看城市、商品类别销售额：使用以下语句查询并得到结果：

文本

描述已自动生成

屏幕上有字

描述已自动生成

### 需求四：

按月查看性别、商品类别销售额：使用以下SQL语句查询并得到如下结果：

文本

描述已自动生成

屏幕上有字

描述已自动生成

## 质量检测：

在对测试的五条用户数据进行质量检测中，使用了电话号码的检测规则，年龄检测规则，邮箱检测规则，得到了如下的质量检测结果：

