**重庆科技学院**

**课程设计任务书**

设计题目：数据导入与预处理实践

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 学生姓名 | |  | | |
| 课程名称 | |  | 专业班级 |  |
| 地 点 | |  | 起止时间 |  |
| 设计内容及要求 | 任务1：对天气观测记录数据的清理，生成适合进行对某地气象预测数据的数据集  基本要求：  1.数据分析  a)数据准确性分析（脏数据：异常值，空值，数据类型，多余数据。。）  b)数据一致性分析（数据量纲、单位，表示法）  c)数据解释性分析（数据的含义）  2.数据处理  a)字段处理方案  b)字段处理详细设计与实现  任务2：根据数据库脚本构建电商业务数据库(itcast\_shop)，通过kettle从业务数据库中抽取数据到数据仓库(itcast\_shop\_学号)中，然后通过sql语句进行数据统计，并将分析结果以图表的形式展示。   1. 数据抽取需求及数据库表结构见附件1 2. 数据分析目标 3. 统计每天订单支付的总金额、订单的总笔数 4. 统计每天下过订单的不同用户总数 5. 统计每天不同支付方式订单总额和订单笔数 6. 统计每天用户订笔数Top5 7. 统计每天不同分类的订单金额和总笔数 8. 统计每天不同商品分类订单的总金额、订单的总笔数   要求：数据抽取过程与统计过程根据需求自动执行。  开发流程如下： | | | |
| 设计  参数 |  | | | |
| 进度  要求 |  | | | |
| 参考资料 |  | | | |
| 其它 |  | | | |
| 说明 |  | | | |

系主任： 指导教师：

年 月 日