# 智慧城市

1. 自我介绍
2. 原来的项目上你主要负责什么工作内容？

数据清洗、转换、写存储过程。

1. Kettle有哪些控件？

表输入、输出、转换、去重、作业等

1. Kettle的控件在工作中是怎么使用的？

数据抽取，作业调度。

1. 作业调度只用kettle吗，有没有其他的方式？

主要是使用kettle，其他的工具用的不多。

1. Kettle有去重的控件？
2. 对数仓的理解？有几层？

数仓主要分为3层，ODS层主要存放业务系统的数据，dw层

1. Sql优化怎么做？

避免使用select \*；

Group by 代替distinct；

尽量多使用COMMIT；

1. 表A两条数据，表B三条数据，用left join关联ID，最后有几条数据？

左关联，2条。

1. 常用的分析函数有哪些？

MAX(),MIN(),SUM(),AVG(),COUNT()

RANK(),DENSE\_RANK()与ROW\_NUMBER()

1. Pg数据库有了解吗？

主要用的是orcle ，不太了解。

1. 存储过程的语法格式
2. 数据是怎样抽取的？
3. 数据库是部署在什么环境的？Windows吗？
4. Linux有用过吗？
5. 不能加时间戳，不能改动原数据的情况下怎么做数据增量更新？

# 环保项目

1. 自我介绍
2. 上一个项目是什么？
3. 工作中主要用什么工具？
4. 除了kettle有了解其他的工具吗？
5. Kettle是什么版本的？
6. 你们数据库是部署在什么环境的？
7. Kettle的组件有哪些？
8. Kettle的增删改用过吗？
9. 你们是在kettle上做作业调度的？kettle不稳定你们怎么解决的，kettle每天都不关闭吗？断电怎么办？

一直在问kettle的问题，说只是使用了一些简单的功能后，面试官表示他们对工具的要求很高，只会简单的跟他们岗位不匹配。就那么愉快的结束了本次面试。

# 软通国际

1. 介绍一下之前的项目流程
2. 主要负责什么内容？
3. 之前工作都哪些事实表和维度表？
4. 建模谁建的，指标是怎么确定的？
5. 怎样确定维度表和事实表要什么样的字段？建好的表最后要得到什么样的结果？
6. DM成的数据是直接做成可视化的报表吗？