**面试总结**

**上海览众科技**

1. 自我介绍和项目介绍
2. ODS层和DW层的定义和区别

第一部分：ODS贴源层一般采用全量的方式从MYSQL同步到ODS贴源层，首先在贴源层建好表框架和结构，主要是有哪些列和数据类型，通过Kettle全量同步数据后，建好转换和作业，再回ODS层检查数据完整性。

区别：数据清洗，过滤空值，日期对不上

第二部分：DW数仓层一般采用维度建模法，根据在ODS层建好的表结构和数据字典进行建表，并且编写存储过程，用Kettle将数据从贴源层同步到数仓层。

区别：进行分类，进行维度和事实表编写存储过程

第三部分：DM集市层根据客户需求编写计算并建表加存储过程，将数据从DW层同步到DM层，再进行可视化操作展示即可。

区别：根据需求做数据模型

1. 为什么想到去上海，上海疫情那么严重都封了。

**上海赛思软件笔试题（全程摄像头看着，只能抄下一些题）**

1. 在表上创建索引后，如何确认索引能被SQL查询所使用
2. Truncate和delete哪个命令更容易产生碎片
3. 请说明以下环境变量的含义

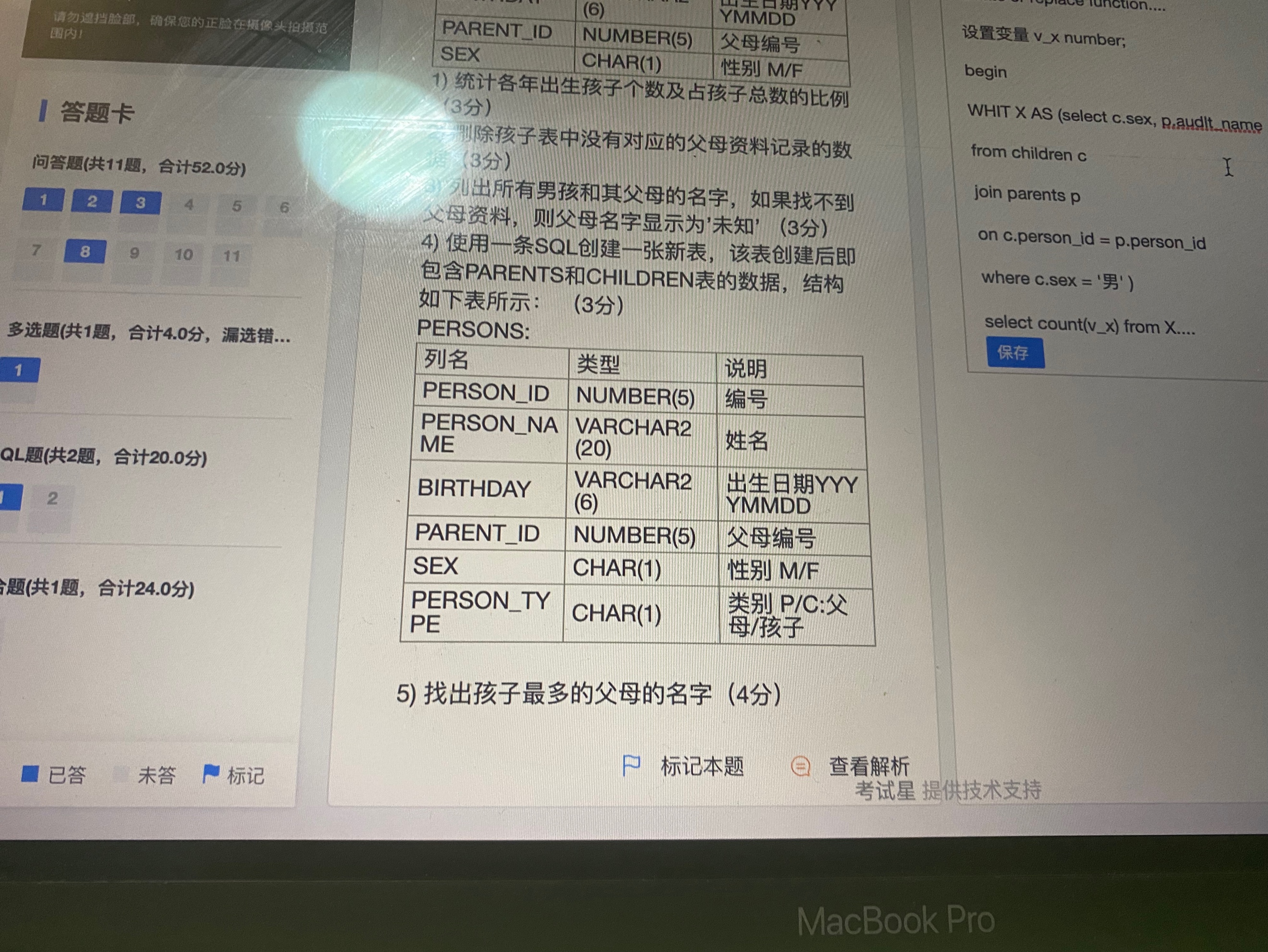
NLS\_LANG

Oracle\_Home

Oracle\_Base

Oracle\_SID

1. 一个OLAP系统Oracle数据库物理内存为72G，数据库参数：sga\_traget = 35G,pga\_aggregate\_traget = 35G,请问此参数有何问题？
2. 业务系统中有以下几张表，customer,product,orderlist,orderdetail,location,画出这些表的E-R，为灵活的分析需求，数据数据集模型，用E-R图来展示出来并且对模型进行说明
3. 举例说明2种数据库中为自增表序号字段赋值的机制
4. 请写出一条sql语句，示范将表中的数据按日期排序，取20-30条记录



**上海中科软**

第一部分面试：

1. 自我介绍
2. 项目介绍
3. 你本科学的主要课程
4. 和国内有什么区别
5. 对未来的规划
6. 是什么原因使你从前一份工作辞职
7. 对未来公司的期望（什么原因使你想留在公司最重要的因素）

第二部分笔试：

用腾讯会议进行桌面共享，做一题讲解一题

第一部分：增删改加建表

第二部分：对题目进行分析并讲解对错（都是比例，计算题）

第三部分：找出数字之间的计算逻辑和规律

**杭州【商智科技有限公司】**

1. 自我介绍
2. 项目介绍
3. 在美国学的课程
4. 专业课和国内有什么不同
5. 讲一下AWS的了解，然后用了哪些api
6. 对数据库的基本理念
7. 对事务的基本特性，ACID有了解过么
8. 对linux了解多少，请说出你知道的linux命令
9. 对索引了解多少，什么时候用到，什么情况下索引会失效
10. 讲对函数计算了解多少，用过哪些函数
11. DW层存储过程的复杂计算讲解一下
12. 这个项目的人员构成，分配的任务
13. 你负责的部分是自己负责么，还是和组员一起帮你
14. 对未来的规划，想做什么

**杭州成功软件笔试题**

**数据处理笔试题**

姓名：\_\_廖健津\_\_ 得分：\_\_\_\_\_\_\_\_\_\_\_

1、请用命令行无密码登陆账户 （可以熟悉的数据库举例，oracle/mysql等）（分值 5）

oracle数据库：

第一种方法：首先以oracle身份登录数据库：su –oracle;然后进入sqlplus控制台，sqlplus/nolog;最后以system的身份登录，命令：connect/as sysdba

第二种方法：在菜单栏里通过改本地ip的方式连接（记不太清了）

2、①什么情况下会造成死锁②发生死锁了可以怎么处理。（可以熟悉的数据库举例，oracle/mysql等）（分值：①5②5）

1. 当某个用户在修改数据或者更新时，会造成资源抢夺的情况，这种时候其他用户不能进行修改，形成死锁现象

2. 但死锁现象发生时，需要得到管理员权限，然后在工具栏中打开会话窗口，找到带锁的会话，点击关掉kill。

1. ①Oracle命令行备份用户scott（密码tiger）的emp表。已知本机能连通数据库服务器，数据库hostname是172.18.0.104，服务名是orcl，域名是18.22.110；

②并还原到本机scott用户下，且本机scott用户下已有名为emp的空表。

（若熟悉的数据库为其他主流数据库，以熟悉的数据库举例也可）（分值：①10；②5）

4、oracle使用什么方法可以使用数据泵备份远程数据库到本地（写用什么技术就可以）。（分值：5，如果能写出具体语句，按正确程度加分1~10）

expdp/impdp命令

5、视图和同义词有什么区别（分值：10）

视图和同义词的区别在于同义词只能对应一张表名称并且可以对其进行任何DML的操作，而视图可以对应多张表，并且只能进行更新和删除的操作。

6、①Oracle的SID、数据库名、全局数据库名、实例名的区别。②安装oracle时设置的默认为“orcl”的这个，是三种里的哪一种。③一般域名是跟哪个一起用的。（分值：①5②2③3）

7、数据库表连接有哪几种，分别举个简单的例子（分值：5）

内连接，左关联，右关联，全外关联，笛卡尔集

8、请写出①查询单字段重复和②查询多字段重复数据的sql语句（分值：①5②10）

用count函数查询数量大于1的重复数据，然后再用group by去重

9、优化下列sql。已知monitortime，pscode已建立索引，且按PARTITION BY RANGE (monitortime) PARTITION P1 VALUES LESS THAN (TO\_DATE('2017-2-1', 'YYYY-MM-DD')方式建立了分区：（分值：20）

SELECT

g.PSCODE, sum(IFNULL(h.so2, 0)) + sum(IFNULL(h.nox, 0)) NUMB

FROM PS\_GAS\_HOUR\_DATA h

INNER JOIN PS\_GAS\_OUTPUT g ON h.outputcode = g.outputcode

where DATE\_FORMAT(h.monitortime, '%Y') = '2018' and g.pscode <> -1

group by g.pscode;

10、什么是ETL？你用过的或接触过的ELT工具有哪些。（分值：5）

ETL: extract, transform,load

ETL工具：kettle