凯美瑞德 招商证券外包

1. 乙方经理初面，聊业务 准备在指标完全不够 硬头皮扯了十几分钟，得到完全不懂业务在评价，感觉马上就要走人了，但是他开始问技术方面，
2. union 和union all 在区别
3. 分析函数用过哪些 做什么用
4. 写过最复杂在逻辑是什么 这个说了一下面试宝典上的 面试官不满意，说还是没有说清楚
5. 大数据系统版本 不知道
6. 大数据做了什么，刚刚开始接触，只做了3个月左右，主要负责从原数据库迁移数据到hive
7. 存储过程在运行过程中，怎么知道它运行到哪里了 没有回答上来 应该是看日志文件
8. 问完后 看他有点纠结，还是带我去找甲方经理面聊，一下来了2个人，觉得阵仗有点大，开始也是问业务，一样是完全不行，他们都感觉庶放弃了，然后开始问技术，
9. 有哪些情况索引会失效 说了 参与运算，对索引字段用了函数（创建基于函数在索引可以解决这个问题），组合函数 如果是只查组合函数的第一是可以走索引 如果是另外的就会失效，他问了在exsits里会不会失效 会
10. SQL怎么写会好些 从表连接开始说 尽量用join 效率最好 最大表放第一位做主表 因为除了主表 其他的都会做全表扫描，也有说法是从下而上，如果有子查询 in 和exsits应用场景有区别 原则是以小表驱动大表 exsits是以外面的表为驱动表，in 是以in里面的为驱动表 然后是书写规范那些 关键字大写 表名T开头，查询中相同的字符大小写保持一致，避免多次硬解析 从这里开始他们才来了兴致，
11. 问一张表有客户、业务、交易额、日期 这个表很大，要统计5年前到当前时间每个客户的交易额的top3的业务，每天会有新数据 要每天更新 表很大 直接top3服务器会崩 想个方案

当场想的方案做一个中间表 存top5 或者top10的数据 这个根据以往的top3变动范围再扩大一些 ，然后再把每天更新的数据合进去，算新一天的，

面试官回复是不一定对 但是有思路

又问了一张用户登录日志表 有用户、登录时间这里直接取日期 sql查询出连续3天都登录的用户 跟之前周老师讲NBA夺冠的一样

 Select user ,count(inum)from

 (Select user,date-rowid as inum from log order by user,date )a

Group by user

Where count(inum)>=3

没有验证 只是思路 面试官认可 结束了 后面说他们这一层算过了

前面业务确实懵 好在技术这块打到他们的点上了