1.最近的一个项目中有哪些亮点

跟他扯自己准备的指标，使劲扯

2.可以独立完成风控相关的建模吗

可以，因为业务很熟了，指标这块计算不会出现太大的偏差

（漏了很多）

3.项目里数仓架构是怎么样的，层次式怎么划分的

这个就完全懵了，扯的啥忘记了，但是没有冷场

4.有接触过数据分析，数据挖掘吗？

有做过，按照时间戳来进行前后对比等等

（没回答到点上）

**简单说：**数据挖掘就是从海量数据中找到隐藏的规则，数据分析一般要分析的目标比较明确。
**主要区别：**

1. “数据分析”的重点是观察数据，而“数据挖掘”的重点是从数据中发现“知识规则”KDD（Knowledge Discover in Database）。
2. “数据分析”得出的结论是人的智力活动结果，而“数据挖掘”得出的结论是机器从学习集（或训练集、样本集）发现的知识规则。
3. “数据分析”得出结论的运用是人的智力活动，而“数据挖掘”发现的知识规则，可以直接应用到预测。
4. “数据分析”不能建立数学模型，需要人工建模，而“数据挖掘”直接完成了数学建模。如传统的控制论建模的本质就是描述输入变量与输出变量之间的函数关系，“数据挖掘”可以通过机器学习自动建立输入与输出的函数关系，根据KDD得出的“规则”，给定一组输入参数，就可以得出一组输出量。
**举个简单的例子：**
* 有一些人总是不及时向电信运营商缴钱，如何发现它们？

**数据分析：**通过对数据的观察，我们发现不及时缴钱人群里的贫困人口占82%。所以结论是收入低的人往往会缴费不及时。结论就需要降低资费。

**数据挖掘：**通过编写好的算法自行发现深层次的原因。原因可能是，家住在五环以外的人，由于环境偏远不及时缴钱。结论就需要多设立一些营业厅或者自助缴费点。

5.OLAP,OLTP的区别

只回答了联机事务处理，联机分析处理

（其他实在想不起来，之前还背过）

问到这里差不多十分钟，问当兵的事情，扯了一会

6.常用linux脚本有哪些

不太熟，最近在看某某某写的hadoop关于分布式系统的书

7.接触过元数据吗？

完全忘记这个概念了，也没有冷场，找个理由带过去了

答案：

8.有做数据清洗的工作吗？

有的，ODC层ETL只做抽取，转换，加载，DW-DM会做清洗，按照不同的维度等等

9.数据监控？

也是回答的按照时间戳，取数逻辑，指标计算，

10.三范式

这个没回答出来

介绍下你们的数据仓库

同步过程中如果停电了，怎么确认数据有没有同步

分析函数有哪些，内置函数有哪些

怎么保存历史数据的

怎么优化

还有些增量，全量的问题

多久进行一次数据更新，用增量还是全量等

数据库内置的数据字典在哪里，怎么找

同义词是什么，作用

如何保证数据的连续性 备份可以分为定期备份和实时备份

执行计划

现场编写九九乘法表

DECLARE

 X NUMBER := 0;

 Y NUMBER := 0;

BEGIN

 WHILE X < 10 LOOP

 X := X + 1;

 DBMS\_OUTPUT.NEW\_LINE();

 --FOR Z IN 1 .. 9 LOOP

 FOR Z IN X .. 9 LOOP

 Y := Z;

 DBMS\_OUTPUT.PUT( X || '\*' || Y || '=' || X \* Y);

 DBMS\_OUTPUT.PUT( ' ' );

 END LOOP; --FOR LOOP

 END LOOP; --WHILE LOOP

END;

还问了几个概念的，资料查不到