# 1：你们是按照哪些主题建模的

客户：主要组织和存放与银行客户有关的信息。包括基本信息、地址信息、信用信息、黑名单信息、财务信息等。在客户主题域中以客户号为唯一识别，通过客户号与存款、贷款、银行卡、中间业务、渠道、公用主题进行关联。

存款；组织和存储企业和个人客户的在银行的存款业务相关信息，主要包括账户信息、事件信息及事故信息。主要分为按个人活期、个人定期、企业活期、企业定期四个子主题。在存款主题域中以账号为唯一识别，通过账号与客户，中间业务、渠道、银行卡、公用主题进行关联。

贷款：组织和存储客户的所有贷款业务数据。根据客户的性质，将贷款客户划分为企业贷款和个人贷款两类。在贷款主题域中以账号为唯一识别，通过账号与客户、公用主题进行关联。

银行卡:组织和存储客户银行卡的基本信息和交易信息。在银行卡主题域中以卡号为唯一识别,通过卡号与存款、客户、渠道、公用。

中间业务：主要整合银行除存、贷款业务以外的业务，即非利息收入以外的所有业务。中间业务主题逻辑划分按中间业务种类进行划分，如国内结算业务、银保通、证券基金、外汇买卖等业务相关信息。在中间业务主题域中以客户号、账号、产品、机构、渠道为唯一识别,分别通过客户号、账号、产品、机构、渠道与客户、存款、公用主题进行关联。

渠道：主要存储渠道信息、签约账户信息、渠道账户信息以及交易流水信息。根据客户性质可将渠道数据分为企业客户和个人客户。在渠道主题域中以客户号、账号、产品、机构、渠道为唯一识别, 分别通过客户号、账号、产品、机构、渠道与客户、存款、银行卡、公用主题进行关联。

总账：组织和存储银行当前会计核算总账以及内部帐有关的信息。在总账主题域中以产品、机构为唯一识别,分别通过产品、机构与公用主题进行关联。

公用：用于存储各种业务主题公用的一些信息。主要整合内部机构、人员、公共代码等相关信息。包括统一标准代码，以及标准代码与各个源系统代码的映射及人工补充代码。

# 2：有贷款项目哪些维度表 事实表



# 3：客户表上有哪些信息

### 客户id 性别 年龄 所在地 职业 学历什么的

# 4：银行数据来源于哪些系统

## 核心系统 表有1000多张

## 客户系统 交易系统 银行卡系统 理财系统 信贷系统 存款系统 网银系统 学习网系统 征信系统 。。。。。一大堆

# 5：计算指标

# 6: 用什么工具抽取数的

## 银行一般用ds 接下来会问用了哪些组件

# 7：历史数据这么保存的，

# 用拉链表，什么是拉链表 有哪些字段

# 8：你们建模原则是，或者是什么思路

* 根据关键业务要素，或业务关注视角，及关键业务要素（业务关注视角）之间的关系，对数据模型进行主题划分；
* 基础数据层的主题划分是通过抽象银行业经营活动中的要素及要素之间关系的形式，来表达商业银行的实际业务和具体的业务联系。它是独立于业务应用需求的，具有高度的稳定性和可扩展性；
* 共性加工数据层的主题划分则是基于业务关注的视角，也就是和业务应用需求紧密相关，会根据应用系统的共性需求变化而变化；
* 不同的数据层次，由于其业务关注视角不同，其主题划分的结果可以不同；
* 各模型层次的主题可根据实际情况划分二级主题，便于用户定位所需数据；

（百度的答案）

# 9：你们数仓规范是什么样的

# 反问，你说的是哪个规范 命名吗 还是文档管理 命名和管理都说说。

# 10大数据

## Hive

## Linux命令

# 11：数据流转你知道

# 12：数据倾斜

# 13：你shell 脚本吗

# 14：Hive 常用函数

# 15：数仓底层原理你知道吗

# 16：你知道外部表和内部表的区别吗

# 17：分区表怎么建立的 分桶

# 18数据治理有了解吗

# Sql 基本没问