您在构建数据仓库之前，首先需要确定构建数据仓库的目标与需求，并进行全面的业务调研。您需要了解真实的业务需求，以及确定数据仓库要解决的问题。

**业务调研**

充分的业务调研和需求分析是数据仓库建设的基石，直接决定数据仓库能否建设成功。在数仓建设项目启动前，您需要请相关的业务人员介绍具体的业务，以便明确各个团队的分析员和运营人员的需求，沉淀出相关文档。

您可以通过调查表和访谈等形式详细了解以下信息：

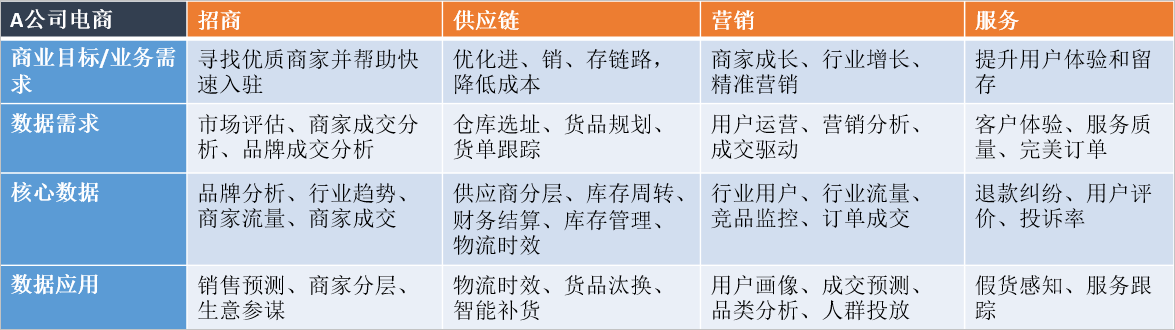
1. 用户的组织架构和分工界面。

例如，用户可能分为数据分析、运营和维护部门人员，各个部门对数据仓库的需求不同，您需要对不同部门分别进行调研。

1. 用户的整体业务架构，各个业务板块之间的联系和信息流动的流程。

您需要梳理出整体的业务数据框架。

1. 各个已有的业务板块的主要功能及获取的数据。

本教程中以A公司的电商业务为例，梳理出业务数据框架如下图所示。A公司的电商业务板块分为招商、供应链、营销和服务四个模块，每个板块的需求和数据应用都不同。您在构建数据仓库之前，首先需要明确构建数据仓库的业务板块和需要具体满足的业务需求。[](https://static-aliyun-doc.oss-cn-hangzhou.aliyuncs.com/assets/img/zh-CN/2680092951/p44519.png)

此外，您还需要进一步了解各业务板块中已有的数据功能模块。数据功能模块通常和业务板块紧耦合，对应一个或多个表，可以作为构建数据仓库的数据源。下表展现的是一个营销业务板块的数据功能模块。

| **数据功能模块** | **A公司电商营销管理** |
| --- | --- |
| 商品管理 | Y |
| 用户管理 | Y |
| 购买流程 | Y |
| 交易订单 | Y |
| 用户反馈 | Y |

**说明** Y代表包含该数据功能模块，N代表不包含。

本教程中，假设用户是电商营销部门的营销数据分析师。数据需求为最近一天某个类目（例如，厨具）商品在各省的销售总额、该类目Top10销售额商品名称和各省客户购买力分布（人均消费额）等，用于营销分析。最终的业务需求是通过营销分析完成该类目的精准营销，提升销售总额。通过业务调研，我们将着力分析**营销**业务板块的**交易订单**数据功能模块。

**需求分析**

在未考虑数据分析师和业务运营人员的数据需求的情况下，单纯根据业务调研结果构建的数据仓库可用性差。完成业务调研后，您需要进一步收集数据使用者的需求，进而对需求进行深度的思考和分析。

需求分析的途径有两种：

* 根据与分析师和业务运营人员的沟通获知需求。
* 对报表系统中现有的报表进行研究分析。

在需求分析阶段，您需要沉淀出业务分析或报表中的指标，以及指标的定义和粒度。粒度可以作为维度的输入。建议您思考下列问题，对后续的数据建模将有巨大的帮助：

* 业务数据是根据什么（维度、粒度）汇总的，衡量标准是什么？例如，成交量是维度，订单数是成交量的度量。
* 明细数据层和汇总数据层应该如何设计？公共维度层该如何设计？是否有公共的指标？
* 数据是否需要冗余或沉淀到汇总数据层中？

举例：数据分析师需要了解A公司电商业务中厨具类目的成交金额。当获知这个需求后，您需要分析：根据什么（维度）汇总、汇总什么（度量）以及汇总的范围多大（粒度）。例如，类目是维度，金额是度量，范围是全表。此外，还需要思考明细数据和汇总数据应该如何设计、是否是公共层的报表及数据是否需要沉淀到汇总表中等因素。

需求调研的分析产出通常是记录原子与派生指标的文档。