**数据库与数据仓库的区别**

1. **从定义角度来看**

数据仓库，是一个面向主题的、集成的、相对稳定的、反映历史变化的数据集合，用于支持管理决策。它是为企业所有级别的决策制定过程，提供所有类型数据支持的战略集合。其出于分析性报告和决策支持目的而创建。为需要业务智能的企业，提供指导业务流程改进、监视时间、成本、质量以及控制。

而数据库是“按照数据结构来组织、存储和管理数据的仓库”。是一个长期存储在计算机内的、有组织的、可共享的、统一管理的大量数据的集合。

一般来说数据库(操作型数据库)是面向事务的，而数据仓库是面对主题的；数据库存储的是业务数据，而数据仓库存储的是历史数据。

1. **从目的角度来看**

数据仓库的主要目的是用来进行数据分析，为企业提供一个决策分析用的工具，帮助决策人员更好地制定企业策略，或找出企业的潜在问题，提高客户满意度，最终提高企业竞争力。

数据库的主要目的是捕获数据并高效且条理分明地存储数据，使人们能够更加迅速和方便地管理数据。便于实现数据共享、减少数据的冗余度、保持数据的独立性、实现数据的集中控制、维持数据一致性和可维护性、易于从故障中恢复。

1. **从模型角度来看**

数据仓库有范式建模方法、维度建模方法和实体建模方法。而维度建模方法应用最为广泛，其中有星型模型、雪花模型、事实星座模型三种模式。

* 星型模型是一种非正规化的结构，每个维度都直接与事实表连接，不存在渐变维度，其有一定的冗余，但有查询效率高、维护方便
* 雪花模型是星型模型的变种，其中某些维表是规范化的，有一个或多个维表没有直接连接到事实表上，而是通过其他维表连接到事实表上。其冗余小，但查询效率低，维护困难。
* 事实星座模型由多个主题构成，包含多个事实表，维表共享。

数据库有分层数据库模型、关系模型、网络模型、面向对象的数据库模型、实体关系模型、文件模型等模型。使用什么模型取决于数据库管理系统是否支持、是否符合设计需求、是否适用于当前设计阶段等等因素。

1. **从设计原则角度来看**

数据仓库模型的设计原则是“自顶向下、逐步求精”。在数据仓库中不需要严格遵守规范化设计原则。其设计分为三个阶段：

* 概念模型，对业务进行系统划分，从整体上进行抽象和概括，也就是主题划分。
* 逻辑模型，进一步细化概念模型。定义实体和实体间的关系，和每个实体的属性。即定义具体表的作用，表与表的约束，表的字段。形成ER图。这些实体的设计都是基于业务规则，可以说，这一阶段主要面对的是业务。
* 物理模型， 具体实现。依照逻辑模型，在数据库中进行建表、索引等。数据仓库，为了满足高性能的需求，可以增加冗余、隐藏表之间的约束等反第三范式操作。

数据库模型的设计原则：

* 一致性原则：对数据来源进行统一、系统的分析与设计，协调好各种数据源，保证数据的一致性和有效性。
* 完整性原则：数据库的完整性是指数据的正确性和相容性。要防止合法用户使用数据库时向数据库加入不合语义的数据。对输入到数据库中的数据要有审核和约束机制。
* 安全性原则：数据库的安全性是指保护数据，防止非法用户使用数据库或合法用户非法使用数据库造成数据泄露、更改或破坏。要有认证和授权机制。
* 可伸缩性与可扩展性原则：数据库结构的设计应充分考虑发展的需要、移植的需要，具有良好的扩展性、伸缩性和适度冗余。
* 规范化原则：数据库的设计应遵循规范化理论。规范化的数据库设计，可以减少数据库插入、删除、修改等操作时的异常和错误，降低数据冗余度等。

数据库的设计中必须严格遵守三范式。

1. **从处理类型角度来看**

数据库属于操作型处理，即联机事务处理 OLTP（On-Line Transaction Processing，），也可以称面向交易的处理系统，它是针对具体业务在数据库联机的日常操作，通常对少数记录进行查询、修改。用户较为关心操作的响应时间、数据的安全性、完整性和并发支持的用户数等问题。传统的数据库系统作为数据管理的主要手段，主要用于操作型处理，像Mysql，Oracle等关系型数据库一般属于OLTP。

数据仓库属于分析型处理，叫联机分析处理 OLAP（On-Line Analytical Processing）一般针对某些主题的历史数据进行分析，支持管理决策。

* 数据库是面向事务的设计，数据仓库是面向主题设计的。
* 数据库一般存储业务数据，数据仓库存储的一般是历史数据。
* 数据库设计是尽量避免冗余，一般针对某一业务应用进行设计，比如一张简单的User表，记录用户名、密码等简单数据即可，符合业务应用，但是不符合分析。数据仓库在设计是有意引入冗余，依照分析需求，分析维度、分析指标进行设计
* 数据库是为捕获数据而设计，数据仓库是为分析数据而设计。