数据库：传统的关系型数据库的主要应用，主要是基本的、日常的事务处理，例如银行交易。

数据仓库：数据仓库系统的主要应用主要是OLAP（On-Line Analytical Processing），支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。

我尝试着再补充些具体的事例来说明，这样更可以帮助大家更好理解一些。

举个最常见的例子，拿电商行业来说好了。

基本每家电商公司都会经历，从只需要业务数据库到要数据仓库的阶段。

* 电商早期启动非常容易，入行门槛低。找个外包团队，做了一个可以下单的网页前端 + 几台服务器 + 一个MySQL，就能开门迎客了。这好比手工作坊时期。
* 第二阶段，流量来了，客户和订单都多起来了，普通查询已经有压力了，这个时候就需要升级架构变成多台服务器和多个业务数据库（量大+分库分表），这个阶段的业务数字和指标还可以勉强从业务数据库里查询。初步进入工业化。
* 第三个阶段，一般需要 3-5 年左右的时间，随着业务指数级的增长，数据量的会陡增，公司角色也开始多了起来，开始有了 CEO、CMO、CIO，大家需要面临的问题越来越复杂，越来越深入。高管们关心的问题，从最初非常粗放的：“昨天的收入是多少”、“上个月的 PV、UV 是多少”，逐渐演化到非常精细化和具体的用户的集群分析，特定用户在某种使用场景中，例如“20~30岁女性用户在过去五年的第一季度化妆品类商品的购买行为与公司进行的促销活动方案之间的关系”。

这类非常具体，且能够对公司决策起到关键性作用的问题，基本很难从业务数据库从调取出来。原因在于：

1. 业务数据库中的数据结构是为了完成交易而设计的，不是为了而查询和分析的便利设计的。
2. 业务数据库大多是读写优化的，即又要读（查看商品信息），也要写（产生订单，完成支付）。因此对于大量数据的读（查询指标，一般是复杂的只读类型查询）是支持不足的。

而怎么解决这个问题，此时我们就需要建立一个数据仓库了，公司也算开始进入信息化阶段了。数据仓库的作用在于：

1. 数据结构为了分析和查询的便利；
2. 只读优化的数据库，即不需要它写入速度多么快，只要做大量数据的复杂查询的速度足够快就行了。

那么在这里前一种业务数据库（读写都优化）的是业务性数据库，后一种是分析性数据库，即数据仓库。

最后总结一下：

数据库 比较流行的有：MySQL, Oracle, SqlServer等  
数据仓库 比较流行的有：AWS Redshift, Greenplum, Hive等

这样把数据从业务性的数据库中提取、加工、导入分析性的数据库就是传统的 ETL 工作。现在也有一些新的方法，这展开说又是另一件事情了，有机会再详细说说。