Data Warehousing

Database Integration

Data Warehousing

OLTP Databases

A database that processes standard business transactions like an ERP or CRM. Tends to involve a mix of reads and writes.
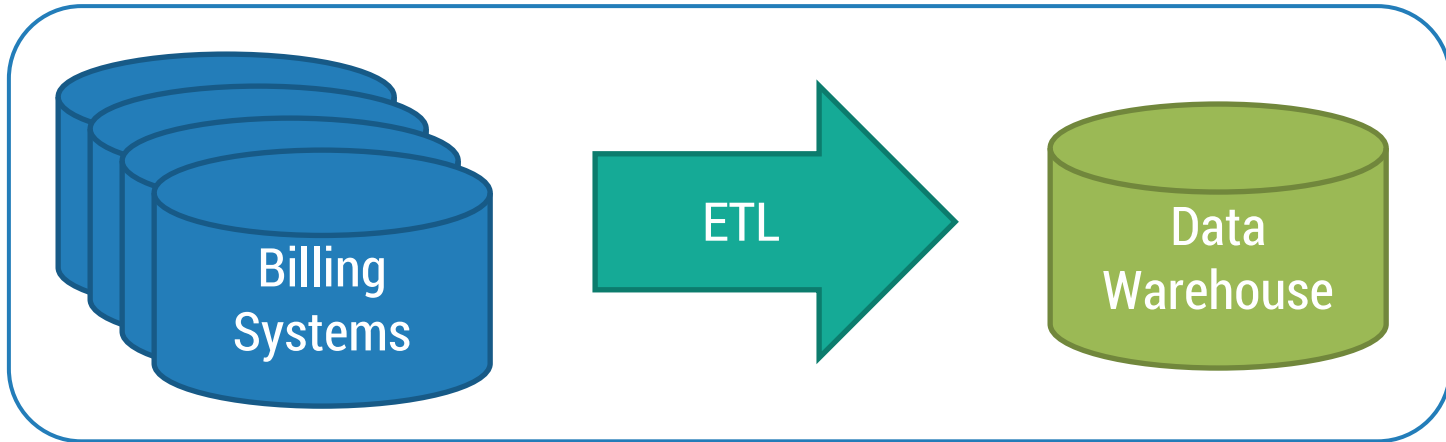
OLAP Databases

Focused on reporting and analytics. OLAP workloads involve database reads over large datasets (like querying the number of orders per day over the last 12 months)
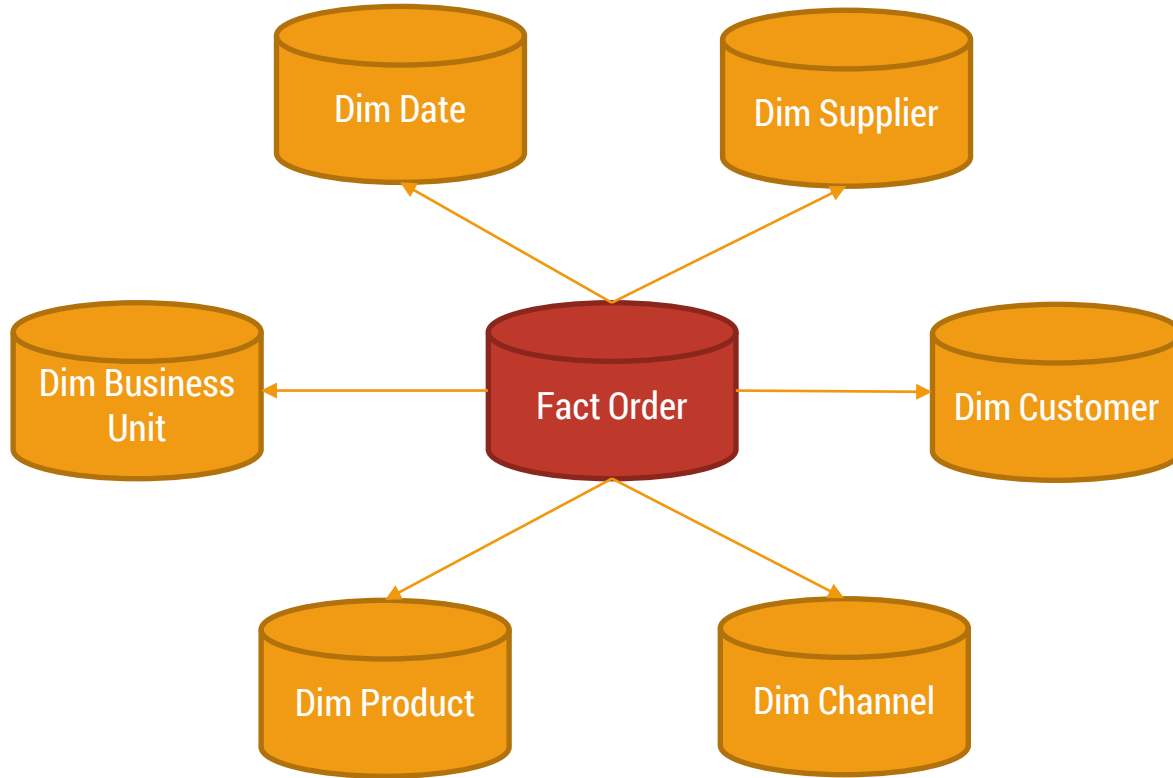
Data Warehouses for Analytics

Annual Revenue Comparison

Channel	2019 Revenue (USD)	2020 Revenue (USD)	Change	%
Direct to Consumer	1,099,912	1,826,741	726,829	66%
Retailers	5,192,881	7,261,812	2,068,931	40%
Distributors	3,299,184	1,928,711	(1,370,473)	-42%
Total	**9,591,977**	**11,017,264**	**1,425,287**	**15%**

Star Schema

Data Lake Architecture

Unstructured Data

Semi-Structured Data

Structured Data

Data Lake
(i.e. Apache Hadoop, S3)

Query Tool (i.e. Map Reduce, Spark)

Data Warehouse