THE DATA LIFECYCLE

The Data Cycle

The Data Cycle

Extraction

Staging Area

Transform & Load

Transform

Data Warehouse

Analytics
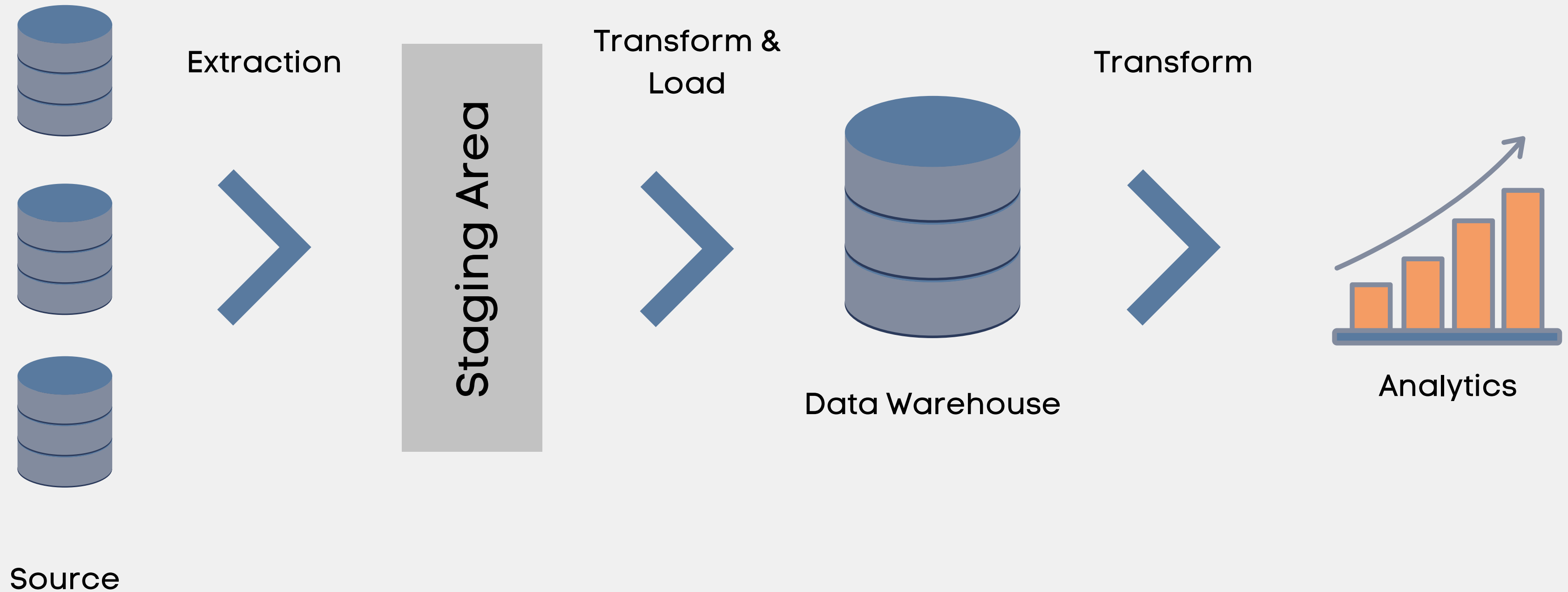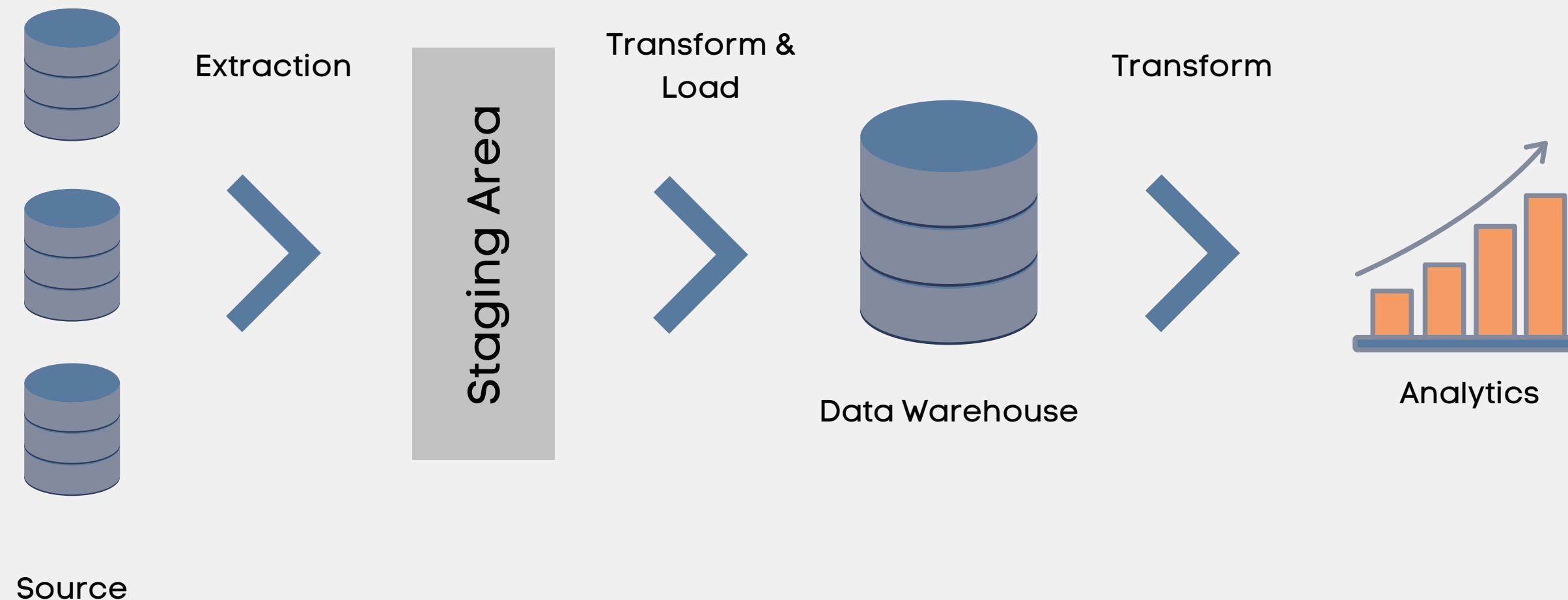
Source

- **Source**: *Data sources may include relational databases and data from SaaS applications. Most pipelines ingest raw data from multiple sources via a push mechanism, an API call, a replication engine that pulls data at regular intervals, or a webhook. Also, the data may be synchronized in real time or at scheduled intervals.*
- **Staging Area:** *A data staging area is a temporary storage area between the data sources and a data warehouse.*
- *The staging area is mainly used to quickly extract data from its data sources, minimizing the impact of the sources.*
- **Transformation and Loading:** *Transformation refers to operations that change data, which may include data standardization, sorting, deduplication, validation, and verification. The ultimate goal is to make it possible to analyze the data.*
- **Data Warehouse:** *A data warehouse (DW) is a digital storage system that connects large amounts of data from different sources to feed BI, reporting, and analytics.*