# What are the pros and cons of using a data warehouse?

There are many pros to using a data warehouse; however, it is imperative that yourself or others in the company are versed within the preparation and implementation of a data warehouse. In addition, this involves entire organization versus single department(s).

|  |  |  |
| --- | --- | --- |
| **Pros** | | **Cons** |
| * Provides support for decision making a , faster business insights b | | * Needs common format for data throughout company, must conform to uniform structures and formats a |
| * Required common format enhances decision making, helping managers better understand company’s operations a | | * Common format integration can be time consuming a |
| * Integrates data derived from entire company, multiple sources, and formats a | | * Time ID assigned to data cannot change a |
| * Centralized, consolidated database a | | * DBMS must support large databases since data is never deleted a |
| * Data is arranged and optimized, providing answers to questions across functional areas in a speedier manner a | | * Requires time, money, and considerable managerial effort a |
| * Not subject to numerous real-time data updates a | | * Interface can be complex for beginners/entry level users b |
| * Represents flow of data through time (historical archive c) or projected data generated through models a | | * Input errors can damage integrity of archived information c |
| * All time-dependent aggregations are recomputed during periodic uploads a | | * Long implementation times |
| * Better data quality b | |  |
| * Supports large-scale BI functions (e.g., data mining, AI, ML)b | |  |
| * Competitive advantage, finding more opportunities in data versus from disparate data stores b | |  |
| a | (Coronel & Morris, 2019) | |
| b | (IBM, 2021) | |
| c | (Frankenfield & Anderson, 2021) | |

# What is the process to move data from the operational warehouse to the data warehouse?

The ETL process is used to move data from the operational warehouse to the data warehouse (Coronel & Morris, 2019). ETL stands for extract, transform, and load. So, the data is extracted (E) from the operational database/warehouse then it goes through a data filter stage (T), and finally gets loaded (L) into the data warehouse. The transformation stage is where the work is done to prepare/format the operational data for the data warehouse. Activities such as filter, transform, integrate, classify, aggregate, and summarize are a part of the transform (T) stage in the ETL process. The graphic on the following page provides a high-level view of the ETL process.

Graphical user interface

Description automatically generated

Source: (Coronel & Morris, 2019)

# Would you recommend the use of a data warehouse and why or why not?

Yes, I would recommend the use of a data warehouse. The preparation/set-up of a data warehouse is intensive and can seem very painful to managers/executive leaders requiring time, money, and considerable managerial effort; however, the value it will provide will make up for that loss. With a data warehouse, a company has a single source of truth for their analytics, and they can make their data work for them. Today, we are inundated with data which can become overwhelming for organizations and people; however, the steps of bringing the different departments together and discussing how data really does flow across the organization will provide individuals with a better understanding of the data/processes as well as help each department understand they are not the only users of the data. As research suggests, data will just keep growing in the coming years….those companies that are taking the time to go through the appropriate steps to implement data warehouses will be ahead of the game and keep that competitive edge.

# What is a data cube and why is it important? How does it work?

## What is a data cube?

A data cube is a three-dimensional cube that represents a conceptual view of multidimensional data (Coronel & Morris, 2019). It is important to note that a data cube is conceptual and NOT how the data is physically stored within the data warehouse. The data cube is made up of facts (e.g., numeric measurements or values like $125,000) and dimensions (e.g., product, time, location). Dimensions provide additional perspective to facts.

Diagram

Description automatically generated

Source: (Coronel & Morris, 2019)

## Why is a data cube important?

Data cubes allow managers/leaders across a company to have different views of the data, resulting in more detailed analysis (Coronel & Morris, 2019). The conceptual three-dimensional model adds perspective to each fact (data value). For example, with sales being the ‘fact’, a product manager may want to view the sales of a product dimension and the store manager may want to see sales by store dimension. Using the data cube, both managers can retrieve only the information they are looking for without adding any additional work. Using time dimensions (e.g., quarter, month, week, day, etc.) provides framework where sales patterns can be analyzed/predicted.

## How does a data cube work?

According to our textbook, “The location of each data value in the data cube is based on its x-, y-, and z-axes.” (Coronel & Morris, 2019) and the three axes signify the dimensions of the data value. For example, a three-dimensional view of sales (data value) by product (y axis), location (z axis), and time (x axis). The actual sales facts are stored in the intersection of each product, time, and location dimension. See figure below for a graphical representation of this view. Lastly, it is important to note that data cubes are considered static, so it must be created before used.

Diagram

Description automatically generated

Source: (Coronel & Morris, 2019)

References

Coronel, C., & Morris, S. (2019). *Database systems: Design, implementation, & management* (13th ed.). MindTap - Cengage Learning.

Frankenfield, J., & Anderson, S. (2021). *Understanding data warehousing.* <https://www.investopedia.com/terms/d/data-warehousing.asp>

IBM. (2021). *What is a data warehouse?* <https://www.ibm.com/cloud/learn/data-warehouse>