Trends and Challenges with Data Warehousing

by Sophia

☰	WHAT'S COVERED

Although businesses understand the role that data driven decision-making plays in their overall effectiveness, many fail to either fully utilize the data they collect, or they make the wrong inferences from the data collected. Business intelligence and analytics can provide businesses and organizations with the information that managers can use to improve the business and increase revenue, thus making the data warehouse an essential piece of any business intelligence strategy. However, while the concept of a data warehouse is simple, it can be very challenging to implement. In this tutorial, we will discuss the challenges and trends associated with data warehousing.

Our discussion will break down as follows:

1. Data Warehouse Challenges

A data warehouse is an extremely large database. For this reason, designing and implementing one is an immense undertaking. The process of design and implementation involves a great deal of planning, collaboration, and coordination of people, resources, and time. As with any large-scale project, there are challenges associated with building a data warehouse. However, if each challenge is addressed properly, the benefits of data warehousing are immediate and lasting for businesses and organizations. The challenges associated with data warehousing are as follows:

- **Data Quality**: In a data warehouse, data is coming from multiple sources within an organization. Data warehouses that include inconsistent data will encounters errors. Inconsistent data, duplicates, and missing data all result in data quality challenges. These quality challenges can result in faulty reporting and analytics necessary for optimal decision-making.

- **Understanding Data**: When building a data warehouse, analytics and reporting will have to be taken into design considerations. In order to do this, the business user will need to know exactly what analysis will be performed. Envisioning these reports will be difficult for someone who has no experience in business-intelligence data analysis, and who is unaware of its capabilities.

- **Testing**: Data must be 100 percent accurate or a business leader could make improper decisions that are detrimental to the future success of their business. This high reliance on data quality makes testing the data warehouse a high priority issue that will require a lot of resources to ensure the information provided is accurate.

- **Performance**: A data warehouse must be carefully designed to meet overall performance requirements. While the final product can be customized to fit the performance needs of the organization, the initial overall design must be carefully thought out to provide a stable foundation from which to start.

- **Design of the Data Warehouse**: Usually, business leaders understand what they need and want out of a data warehouse. However, if they don't fully understand all the implications of these needs and wants, they will have a difficult time adequately defining them. Often this results in miscommunication between the business users and the technicians building the data warehouse. The typical end result is a data warehouse which does not deliver the results expected by the user. Since the data warehouse is inadequate for the end user, there is a need for fixes and improvements immediately after initial delivery. The unfortunate outcome is greatly-increased development costs.

- **Cost**: There are a multitude of hidden problems in building data warehouses. A frequent misconception among business leaders is that they can build a data warehouse internally to save money. Even if a business or organization adds a data warehouse "expert" to their staff, the depth and breadth of skills needed to deliver an effective result is simply not feasible with one or a few experienced professionals leading a team of non-business intelligence trained technicians.

- **Privacy Concerns**: The increasing power of data mining has caused concerns for many, especially in the area of privacy. In today's digital world, it is becoming easier than ever to take data from disparate sources and combine them to do new forms of analysis. In fact, a whole industry has sprung up around this technology: data brokers. These firms combine publicly-accessible data with information obtained from the government and other sources to create vast warehouses of data about people and companies that they can then sell.

2. Data Warehousing Trends

Data warehousing is no longer an abstract idea; it is a reality. A significant number of businesses and organizations across the globe have made the commitment to building and maintaining data warehouses. Data warehousing has made an enormous impact on the way people perform business analysis and make strategic decisions. Companies that incorporate data warehouses realize benefits that positively affect their bottom line. As data warehousing will only increase due to more and more companies realizing its effectiveness, the major trend seems to be pointing toward the web-enabled data warehouse. A web-enabled data warehouse will involve the real time of capturing the screenshots and clickstream (mouse clicks) of visitors to a business or organization's website, in addition to performing all of the traditional data warehousing tasks. Adding this layer of functionality to the traditional data warehouse will enable businesses and organizations to analyze: web traffic, effectiveness of marketing campaigns, customer satisfaction, demographic data collection, user preferences, affiliate product relationships, customer buying patterns, and website feedback.

> ✅ SUMMARY
>
> Seeking to gain a competitive advantage, businesses and organizations have embraced the idea of data warehousing as a central part of their **business intelligence** strategy. In this tutorial, we took a deeper look into the **data warehouse** to assess the challenges associated with **building a data warehouse**, as well as the trends that are informing the future of data collection amongst businesses and organizations.

Source: Derived from Chapter 4 of "Information Systems for Business and Beyond" by David T. Bourgeois. Some sections removed for brevity.

https://www.saylor.org/site/textbooks/Information%20Systems%20for%20Business%20and%20Beyond/Textbook.html