

Original

If your business has outgrown the reporting that comes with your business apps, you may be looking at implementing a data warehouse. But where to start? The technology can be confusing with all the options available.

We get this question a lot, so we wrote this guide to help clients understand the basics of what is required for a successful data warehouse. We've broken it out into two major areas - Implementation and Support – because both are critical to a successful project. The good news is that there are only three ingredients you need for both: Hardware, Software, and Expertise.

Implementation refers to the design and initial build of a data warehouse. First up, you need experts who have had experience creating data warehouses for other clients. While both large and small companies may use data warehouses, hiring experts who have done both are valuable as they will know how to scope the specifics of what your business needs. This way, your company only implements the technology required for today and the near future (two to three years) rather than what would be required in a company ten times your size.

Today's hardware requirements have been made much easier by the broad adoption of cloud technologies. Since breaking out the now virtual components of processing power, memory, and storage means they can all be adjusted to fit your exact needs, which results in an easy expansion path for the future.

As you might expect, the software is where it comes together. The essential item you are looking for is a software platform that uses a specific database designed for On-Line Analytical Processing. This "OLAP" database is what you need when building a system for analytics, which differs from the typical type of transactional database that lives under the hood of most business applications. Systems such as an Enterprise Resource Planning (ERP, accounting platform, or Customer Relationship Management (CRM) applications need an On-Line Transaction Processing or "OLTP" database. Designed for processing large volumes from multiple sources, OLTP databases are what you want for those functions.

At this point, you may be wondering why these two types of databases can't just talk to each other to work better together. Sounds simple right? That's where the experts come in. The core

function of a data warehouse is to ensure the data from your OLTP-based applications can get to your OLAP database in the most efficient way possible. Here's the flow (insert graphic)

- ➔ Point of Sale system - *OLTP*
- ➔ Finance/ERP system - *OLTP*
- ➔ Inventory Management system- *OLTP*
- ➔ Data Warehouse
- ➔ Reports & Analytics

To make sure everyone can rely on these reports to make

A word on Data Lakes

One approach that has been used to gather data together in one place is the concept of a "Data Lake". Simply put, it's a repository for both structured and unstructured data that can scale. Azure Data Lake Storage, Amazon S3, Google Cloud Storage or even Hadoop based systems can all fit the definition of a data lake in certain cases. While the benefits to using a Data Lake may seem tempting, the long-term effect that it has on an analytics team is substantial.

The main selling feature of "dump all your data, structured or not, and make it easy" works against the analysts who are attempting to make sense of it as they have to create organization where none exists. The storage layer is really cheap, and easily accessible by technology but dramatically increases the workload on the consumer to interpret it. By comparison, a data warehouse or even a "lake house" applies structure to the data before any analytic need arises. Over the long-term, as additional analytics are requested or become more complex, re-doing the work interpreting the data becomes too onerous and the data lakes stagnate into "data swamps". Since good analysis is all about repeatability and consistency, the ongoing reinterpretation required by data lakes makes them a poor stand-alone solution.

[A Quick Introduction to Data Warehousing](#)

As your business outgrows the ~~built-in~~ reports ~~built into~~ of your ~~operational~~ programs, you may look at ~~implementing~~ a data warehouse ~~to be a foundation~~ for analytic reporting. But ~~the technology all the technology options available~~ can be confusing ~~with all the options available~~.

Where to start?

We get this question a lot, so we wrote this guide to help clients understand the basics. We've focused on the two major areas [critical to a successful data warehouse project](#)—implementation and support—~~critical to a successful data warehouse project~~, which share three vital, overlapping requirements: hardware, software, and expertise.

[Plan Your Implementation](#)

Implementation is the design and initial build of a data warehouse before installation. ~~The first~~[Each](#) decisions ~~made now~~ will affect every other step in the process, so you should start with experienced guidance.

[Hire Experts](#)

While ~~both~~ large and small [\(and in-between\)](#) companies may use data warehouses, they don't always need the same functionality. Experts who have built data warehouses for companies of various sizes know how to scope the specifics of what your business needs.

Measuring the value of great analytics to each organization is a complex question. Design decisions change as technology changes, but the output remains the same: understanding your business ~~efficiency~~, your competition ~~on~~[ve landscape](#), and your opportunities ~~for growth~~.

A well-designed data warehouse ~~can have~~[has](#) many different dimensions:

- [Technology \(how up to date and how powerful\)](#)
- Uptime (how long the system can run before an outage)
- ~~Technology level (how up to date)~~
- ~~Implementation & Operation c~~[Cost](#) (how much to build, [operate](#), and support)

- ~~Business coverage~~[Coverage](#) (~~—how many people use it, how often, across how much of your organization~~)
- [Accessibility \(Ease of Access–Useability \(How how](#) easy the analytics are to access ~~for everyone in your organization)~~

Because technology advances so quickly, with expert guidance your company only invests in ~~the technology~~[what you](#) required for today and the near future (two to three years) rather than wasting resources designing a warehouse for a company ten times your size.

[Select Hardware](#)

Today's hardware requirements for processing power, memory, and storage are much easier to anticipate and manage because of the broad adoption of cloud technologies. Companies no longer *must* commit to physical hardware that may limit ability or become obsolete before reaching its full potential. Instead, experts can select and adjust virtual components individually to fit your exact needs, which results in savings now and an easy expansion path for the future.

[A Word of Warning on Data Lakes](#)

One approach to cloud storage is a “data lake,” a scalable repository ~~with the main selling feature being~~[often bought for the ability](#) to store all your data, structured or not, and make it easy for technologies to access. By comparison, a data warehouse or even a “lake house” applies structure to the data before any analytic need arises.

While a data lake may seem tempting, the long-term effect ~~that~~ it has on an analytics team is substantial; the lack of structure dramatically increases the workload of the analysts who have to create organization where none exists. Over ~~the long-term~~[time](#) as analytics are added or become more complex, constantly re-interpreting the data becomes onerous and costly, ~~while and~~ the data lake stagnates into a “data swamp.” Since good analysis is all about repeatability and consistency, data lakes make a poor stand-alone solution.

Final

A Quick Introduction to Data Warehousing

As your business outgrows the reports built into your programs, you may look at a data warehouse for analytic reporting. But all the technology options available can be confusing. Where to start?

We get this question a lot, so we wrote this guide to help clients understand the basics. We've focused on the two major areas critical to a successful data warehouse project—implementation and support—which share three vital, overlapping requirements: hardware, software, and expertise.

Plan Your Implementation

Implementation is the design and initial build of a data warehouse before installation. Each decision now will affect every other step in the process, so you should start with experienced guidance.

Hire Experts

While large and small (and in-between) companies may use data warehouses, they don't always need the same functionality. Experts who have built data warehouses for companies of various sizes know how to scope the specifics of what your business needs.

Measuring the value of great analytics to each organization is a complex question. Design decisions change as technology changes, but the output remains the same: understanding your business, your competition, and your opportunities.

A well-designed data warehouse has many different dimensions:

- Technology (how up to date and how powerful)
- Uptime (how long the system can run before an outage)
- Cost (how much to build, operate, and support)
- Coverage (how many people use it, how often, across how much of your organization)
- Useability (how easy the analytics are to access)

Because technology advances so quickly, with expert guidance your company only invests in what you require for today and the near future (two to three years) rather than wasting resources designing a warehouse for a company ten times your size.

Select Hardware

Today's hardware requirements for processing power, memory, and storage are much easier to anticipate and manage because of the broad adoption of cloud technologies. Companies no longer *must* commit to physical hardware that may limit ability or become obsolete before reaching its full potential. Instead, experts can select and adjust virtual components individually to fit your exact needs, which results in savings now and an easy expansion path for the future.

A Word of Warning on Data Lakes

One approach to cloud storage is a "data lake," a scalable repository often bought for the ability to store all your data, structured or not, and make it easy for technologies to access. By comparison, a data warehouse or even a "lake house" applies structure to the data before any analytic need arises.

While a data lake may seem tempting, the long-term effect it has on an analytics team is substantial; the lack of structure dramatically increases the workload of the analysts who have to create organization where none exists. Over time as analytics are added or become more complex, constantly re-interpreting the data becomes onerous and costly, and the data lake stagnates into a "data swamp." Since good analysis is all about repeatability and consistency, data lakes make a poor stand-alone solution.

Create Software

As you might expect, software is how data from all your operational programs comes together into analytical reporting.

The essential item when building an analytics system is a software platform that uses a database designed for On-Line Analytical Processing (OLAP). An OLAP database contains dynamic functions for how the system should analyze data.