

Data Mining / Warehousing Best Practices

1. **Make sure you understand the reasons for implementing data warehousing.** If senior management wants to simply know mundane information, such as yesterday's sales figures, it may not make sense. However, if you need to find out when a customer segment ordered certain products, what channels they used, how orders were affected by promotional activities, email campaigns and customer support calls, then it may make sense.
2. **Define your goals and develop a supporting business strategy.** What are your goals in implementing data warehousing? What questions do you need answers for? (For example: Why are sales improving? Which products? Who should I try to sell more to? Through what channel?) Think about how you would manage your business differently if you had this information. This will help you define your goals to meet your business strategies and determine if data warehousing would help.
3. **Determine if you have, or will have, data to support a data warehouse.** Having eCommerce product, order, and customer data is usually not enough to justify the expense of a data warehouse project. You would also need cross-channel data from your stores and customer service organization, in addition to your eCommerce data and analytics, to fully understand your customers and their buying patterns. You will need to have detailed customer profiles/groups, an order history, and full website analytic information.
4. **Identify possible data sources.** Data may come from many sources: online analytics, eCommerce databases, customer support logs, social media, mobile, brick and mortar POS, etc.. Understanding the sources of your data will help you know what integration capabilities the data mining/warehousing solution needs to have to support your environment.
5. **Investigate various data warehouse solutions.** Each solution and architecture has its own set of features and limitations. Understanding their capabilities will help determine how to use these tools and what data you need to collect to support them. Do your own diligence and do not let a vendor identify your needs.
6. **Define the security that you need for your data.** Data security is critical to eCommerce sites. Think about what security needs to be in place to get access to the data. Does it need to be encrypted, behind a firewall, backed up offsite, etc? Does it meet PCI and country/state legislation compliance?
7. **Try to implement in phases.** Implementing your entire solution at once may take too long and cause you to lose focus. Identify and prioritize your top needs, determine what is needed to complete each, and keep phase down to 2-3 months rather than longer than 6 months. Remember, a quick, small win is better than a larger goal that may never be achieved.
8. **Determine your data management strategy.** As your business use of data grows and changes, it is very easy to find yourself in a situation where your data is unusable. Continually cleaning and verifying is much easier than undergoing a massive cleanup project in a year. In addition, as you use it, you will find that your needs change over time. To avoid massive changes, monitor your needs and adjust on an ongoing basis.
9. **Assign the proper resources.** Managing a data warehouse can require at least one full-time position. Tasks such as reporting, maintenance, validation are ongoing. Having a full-time manager may make sense.
10. **Look at the entire cost.** Since some of your needed data sources may not exist currently, you may need to purchase additional products, like web analytics, to support your goals. If you start adding up these additional applications, costs and timelines, your total cost may be much higher than planned.