

Data Mining/Warehousing Best Practices

1. **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. If you need to determine when a customer segment ordered certain products, what channels they used, the effect of promotions, and customer support calls, then it may make sense. If the need involves predictive analytics, then data mining makes sense too.
2. **Define your goals and develop a business strategy.** What are your data warehousing goals? What questions do you need answers for? (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.
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 to fully understand your customers and buying patterns. You will need detailed customer profiles/groups, order history, and full site analytics.
4. **Identify possible data sources.** Data may come from many sources: online analytics, ecommerce databases, customer support logs, social media, mobile, POS, etc. Understanding data sources will help determine integration capabilities.
5. **Investigate various data warehouse solutions.** Each solution has its own features and limitations. Understanding capabilities will help determine how to use these tools and what data you need to collect. Do your own diligence and do not let a vendor identify your needs.
6. **Identify data hygiene issues.** The old phrase "garbage in, garbage out" applies to data warehouse projects. Each data source has its own data cleanliness issues that would need to be addressed. There is also the issue of data format incompatibilities between systems. It might be worth investigating "Master Data Management" solutions if these data hygiene issues are severe enough to cripple efforts toward data normalization.
7. **Determine your data management strategy.** As your use of data changes, you may 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. To avoid massive changes, monitor your needs and adjust on an ongoing basis.
8. **Define your data security needs.** Data security is critical. 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?
9. **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 phases down to 2-3 months. Remember, a quick, small win is better than a larger goal that may never be achieved.
10. **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.
11. **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. As you add up these additional applications, costs and timelines, your total cost may be much higher than planned.

**For 1000s of Best Practices and Advice across 100s of topics:
Join eCommerceKnow-How.com!**

eCommerce Diligence™ Checklist

Data Mining/Warehousing Key Questions to Ask Providers

Company

1. How long have you been in this business?
2. How many clients have you sold? How many have you lost, and why?
3. What size or types of clients fall into your “sweet spot”?
4. What peripheral or supporting services do you offer (e.g., customer LTV analysis)?

Products/Services

1. Is your solution offered as perpetual licensed software or on-demand multi-tenant solution?
 - a. What are the client support requirements for each?
 - b. What is the process to add features you currently don't have? Can I make changes or customize the solution?
 - c. What redundancy and disaster recovery do you have in place?
2. Please describe your architecture.
3. Please describe how your solution scales. What are the levels? Do you provide tools to measure and determine scalability and build out requirements?
4. What is your solution built on? Which BI, OS and database, integration protocols, API's, etc.?
5. What skills does my organization (or hired third party) need to implement this platform?
6. What type of training do you provide?
7. What technical support services are available?
 - a. Are there human beings I can reach during reasonable hours (chat or phone) or do I have to wait over 24 hours for an email response or callback?
 - b. Is your support team located in the US or abroad?
 - c. What is your SLA for support issues? System service? Availability? Performance? Backups?
8. Do you have a support knowledgebase, community forum, or applications that are shared by customers?
9. What are the data destruction and retention policies?
10. How do I ensure I own and control my data?

Features

2. What key features are included in your solution?
 - a. Is there a key feature you have that no one else has?
 - b. What features are currently missing that many of your competitors have, are on your roadmap, do management love?
3. What is your product's competitive advantage over other data warehousing packages? Why?
4. Do you require all the data be synchronized to your data warehouse or can your system pull data from several external systems to generate my reports? Please explain.
5. Please provide a list of external integrations that your solution supports. Indicate whether these are real time or batch, import and/or export capable, and the rules provided to support. Please explain in as much detail as possible.
 - a. What type of connectivity you provide to perform importing and exporting of data (APIs, FTPs, ODBC, etc.)?
6. Please provide a list of supported protocols to obtain or send data to the data warehouse.
7. Does your solution provide a staging area to review and approve data prior to uploading to the live system? Please describe.
8. Please explain your security model and approval process for obtaining access to data, access to specific reports, and modification of reports or the data structure.
9. Please describe your backup and recovery, disaster contingency plans, and SLA to restore normal operating conditions.
10. Do you have a management console? What tools does it include, who can use it? Training?
11. Please provide an explanation on the report generation capabilities, ad hoc reporting, training needed, etc.
 - a. How do you provide reporting? Batch, real time, what latency?
 - b. Please describe your reporting. Please provide a list of standard reports.
 - c. Please provide a list of 3rd party BI solutions you have integrated with. Past, present and future.
12. What tools do you provide to help retailers? (Standard tools to help optimize prices, forecast demand, range analysis, etc.

Pricing

1. Please describe your pricing model.
2. Do you charge for number of users, database size, etc.?
3. What tools are included in the price? Which modules require add-on costs?
4. Are there different levels?
5. Why do I want to pay to upgrade to the next level?
6. How do you charge for annual support? What about maintenance? Upgrades? Feature changes?
7. Are there any additional fees (e.g., implementation, transaction costs, revenue sharing)?

For 1000s of Best Practices and Advice across 100s of topics:
Join eCommerceKnow-How.com!

©FitForCommerce

eCommerce
KNOW-HOW.com™
by Fit forCommerce