CHAPTER 30

RETRIEVING DATA FROM OLAP SERVERS

In this chapter

What Is OLAP? 790
Server Versus Client OLAP 791
Creating an OLAP Data Source Definition 794
Creating an OLAP PivotTable 794
Using OLAP PivotTables 795
Using OLAP PivotCharts 798
Saving Offline Cubes from Server Cubes 799
Performing OLAP Analysis on Database Data 801
WHAT IS OLAP?

Excel 2003’s On-line Analytical Processing (OLAP) features enable you to quickly sift through large amounts of information to get the answers you need to solve business problems and make well-informed decisions. OLAP (pronounced (OH-lap) techniques are especially suited to help you see long-term trends and discover hidden relationships and causal factors in business performance.

OLAP views the world in two categories, measures and dimensions. Measures represent things that you can count or measure in numeric terms. The number of units sold, sales revenues, gallons of fuel, pounds of cargo, and number of employees are examples of typical measures. Dimensions are nonnumeric things that give relevance to measures. A salesman’s name, date of sale, location of sale, and type of product sold are examples of dimensions that describe a sales revenue measure.

Dimensions typically are organized into hierarchies. For example, sales for all the sales persons at a store summarize to the store total; all the stores in a sales district summarize to the district total; all the districts in a sales region summarize to the region total; and all sales regions summarize to the corporate total. In this example, measures are aggregated at each level in the sales organization dimension hierarchy. Another example of a dimension hierarchy is a time dimension where daily sales summarize to a monthly total, monthly sales summarize to a quarterly total, and quarterly sales summarize to an annual total as shown in Figure 30.1.

Figure 30.1
Measures are aggregated at each level of detail in a dimensional hierarchy.

With OLAP, you are freed from the drudgery of entering values into your worksheet. Excel retrieves data values from a database at your request and inserts them into PivotTables. Measures are placed in the inner cells of the PivotTable, and dimensions are placed in the column and row headers, as shown in Figure 30.2. Excel maintains a connection to the data
source (hence the term online) as you rearrange and requery the data through a series of exploratory views in an interactive dialogue with the computer. You may choose to change the dimensions displayed in the PivotTable’s header rows or the measure displayed in the data cells, or you may drill down within a dimension to explode the view to the next level of detail within the dimension hierarchy. Excel quickly refreshes the PivotTable with the new values from the data source. Each view yields information from a different perspective and influences the path you take to create the next view.

**Server Versus Client OLAP**

OLAP depends on a special database structure called a *multidimensional cube* to provide fast response when you drill and change dimensions and measures. The cube organizes the data into dimensions and precalculates the aggregates at each of the dimension hierarchy levels. The decisions about which measures, dimensions, and dimension levels to include in a cube when it is created limit what you can view in an Excel OLAP PivotTable.

With Excel, you can choose to build your own cubes from a relational database, such as Microsoft Access or Oracle, or you can connect to cubes prebuilt by a database administrator and stored in Microsoft SQL Server version 7 or later (see Figure 30.3). With server-based cubes, you can connect to the server and immediately begin pivoting and drilling through cube data. With self-built cubes, first you must run Excel’s wizard to define the cube structure and then wait for Excel to process data from the source into the cube. You then can browse the cube’s data in an OLAP PivotTable and save the cube to disk for later reuse in PivotTables. Because cubes contain snapshots of the source data from which they
are built, whenever data in the underlying source database changes, the cube must be reprocessed to include the new or changed data.

Processing cubes can be very time consuming. Smaller cubes can be built in a matter of minutes, but if the source data has millions of rows or the cube has many measures and dimensions with many hierarchy levels, loading and calculating the data in the cube could take hours. Larger, more complex cubes are best left to SQL Server database administrators who have powerful multiprocessor servers and scheduling software at their disposal. They can automate cube update processing and manage cube refresh schedules to ensure server-based cubes are always up to date. The size and complexity of the cubes you build yourself with Excel will be limited by the power of your personal workstation and your patience.

**Creating an OLAP Data Source Definition**

Microsoft provides the offline database drivers with Office 2003 to connect to SQL Server OLAP cubes. The first step in connecting to cubes is to define a data source connection. After it is defined, you can reuse data source connections in future Excel sessions. You must define a data source connection for each OLAP source you intend to use in Excel. To do so, follow these steps:

1. Choose Data, Import External Data, New Database Query to open the Choose Data Source dialog box shown in Figure 30.4.
2. Click the OLAP Cubes tab, select <New Data Source> in the list, and click OK to open the Create New Data Source dialog box (see Figure 30.5).

Figure 30.5
Give your OLAP data source any name you want, and then select which type of OLAP server you want to use.

3. Type a name for your new data source. In the future, you will select this name from the Choose Data Source dialog box to connect to the data source.

4. Select an OLAP provider from the drop-down list. To connect to Microsoft SQL Server 2000 OLAP cubes, select Microsoft OLE DB Provider for OLAP Services 8.0.

5. Click the Connect button to open the Multidimensional Connection dialog box (see Figure 30.6).

Figure 30.6
The Analysis Server option accesses cubes in Microsoft SQL Server 2000; the Cube File option accesses offline cubes stored on your hard disk.

6. Click the Analysis Server radio button. Type the server name, your user ID, and password in the spaces provided. Click Next.

7. Select the OLAP database from the list of available databases and click Finish to return to the Create New Data Source dialog box.

8. Select the desired cube from the drop-down list of cubes in the database you connected to in the previous step. Select the check box at the bottom of the Create New Data
Source dialog box if you want to save your user ID and password as part of the data source definition—keep in mind that this might pose some security risk.

9. Click OK to save the OLAP data source definition and return to the Choose Data Source dialog box.

You could immediately start to use your new data source by selecting it in the Choose Data Source dialog box and clicking OK, but you learn another path to using OLAP data sources in the following section. Click Cancel to close the Choose Data Source dialog box.

CREATING AN OLAP PIVOT TABLE

To retrieve data from an OLAP cube, you create an OLAP PivotTable in your Excel worksheet. Follow these steps:

1. Choose Data, PivotTable and PivotChart Report to start the PivotTable and PivotChart Wizard (see Figure 30.7).

2. Click the External Data Source and PivotTable radio buttons, and click Next.

3. Click the Get Data button in Step 2 of the wizard (see Figure 30.8) to open the Choose Data Source dialog box.

4. Click the OLAP Cubes tab and select the data source you created previously.

5. Click OK to return to Step 2 of the wizard, and click Next to proceed to Step 3 of the wizard.
6. Choose to put the PivotTable in your current worksheet or in a new worksheet the wizard will create for you (see Figure 30.9). Click Finish to create the OLAP PivotTable.

**Figure 30.9**
Choose whether you want your new OLAP PivotTable to go into its own worksheet or into the one you’re currently using. Click Finish to create the PivotTable.

**USING OLAP PIVOT TABLES**

Excel’s OLAP PivotTable contains four drop areas for you to place data elements (see Figure 30.10). The center drop zone, labeled Drop Data Items Here, is reserved for placing measure fields from your cube. You place dimension fields in the Page, Row, and Column drop areas. While working with OLAP PivotTables, Excel also displays a dockable PivotTable toolbar and a PivotTable Field List. If the PivotTable Field List is not visible, you can click the Show Field List button on the PivotTable toolbar.

**Figure 30.10**
The empty PivotTable contains drop areas for you to place fields from the OLAP data source. You select measure and dimension fields from the PivotTable Field List.
To have a meaningful view, you must place at least one measure field in the Data Area and one dimension field in either the Page, Row, or Column area. You do not have to use all three dimension drop areas. To place a field in the PivotTable, select the field in the PivotTable Field List and use the mouse to drag and drop it on the desired drop area. You also can select the desired drop area from the drop-down list at the bottom of the PivotTable Field List window and click the Add To button to place the selected field in the PivotTable. Excel will retrieve the selected data elements from the cube and refresh the PivotTable view.

The numbers displayed in the Data Area are those values that occur at the intersection of the dimensions displayed on the PivotTable. You can trace a cell value back to its row and column headers to see which dimensions are relevant to that data cell. Note that, although multiple dimension values can be displayed across multiple row and column headers, only a single value can be displayed for each dimension in the Page Area. You can think of the Page Area as a data filter. All the data values in the cells of the Data Area are filtered to match whatever dimension values are displayed in the Page Area.

You can place more than one measure field in the Data Area and more than one dimension field in each of the Page, Row, or Column Areas. Excel will nest the display of multiple fields within a single drop area, as shown in Figure 30.11. Each dimension or measure placed in a drop area displays a shaded PivotTable button. After you place a field on the PivotTable, you can drag and drop PivotTable buttons between the drop areas to rearrange the view. Right-click a PivotTable button, and choose Hide Dimension from the pop-up menu to remove a dimension from the PivotTable. Dimensions can also be removed from the PivotTable by dragging the PivotTable button outside the PivotTable until the mouse pointer displays an “X.”

To drill down in a dimension hierarchy, double-click a dimension header cell in the Row or Column Area. If a lower-level hierarchy exists below the dimension level, Excel will explode the PivotTable view to show the details of the lower level. If the exploded level is already visible, Excel will collapse the view and remove the lower-level detail from the PivotTable. For example, double-clicking Drink in the PivotTable in Figure 30.11 expands the view to show the drink categories Alcoholic Beverages, Beverages, and Dairy, as shown in Figure 30.12. Double-clicking Drink in Figure 30.12 will return to the original view in Figure 30.11. You can also click the Show and Hide Detail buttons on the PivotTable toolbar to expand and collapse dimension hierarchies.

Click the drop-down arrow on a dimension PivotTable button to open a tree view of the dimension, as shown in Figure 30.13. Click the plus and minus signs to expand and collapse
Using OLAP PivotTables

branches of the dimension hierarchy in the tree. You can pick and choose which elements of the dimension to display in the PivotTable by clicking to toggle a check mark on or off in the box to the left of the element. Elements with a check mark will display. Elements without the check mark will be hidden.

Figure 30.11
You can add multiple dimensions to the same drop area. Here the Store Type dimension is nested within the Product Family dimension in the Row Area of the OLAP PivotTable.

Figure 30.12
You expand and collapse the level of detail in a PivotTable by drilling into dimension hierarchies. Double-click a dimension header cell or use the Hide and Show Detail buttons on the PivotTable toolbar to drill up and down hierarchies.
Using OLAP PivotCharts

A PivotTable full of numbers is a good way to see details but is not always the best way to see the big picture. Excel charts are a great way to get an overview of your data at a glance and to reveal hidden information. Pie charts, bar charts, and area charts are good for demonstrating proportions and contributions to the whole of individual dimension elements. Bar charts and line charts are good for revealing trends, particularly across a time dimension. Excel PivotCharts complement PivotTables to provide advanced visual analysis of OLAP data.

NOTE

For more information on charts, see Chapters 20–23. For more on PivotTables, see Chapter 24.

TIP FROM

John

When you hide items, they will not be displayed in the total lines of the PivotTable. Choose the Include Hidden Items in Totals button on the PivotTable toolbar if you want the totals to reflect hidden values.
SAVING OFFLINE CUBES FROM SERVER CUBES

A PivotChart has the same Data Area and Page Area as the PivotTable, but maps the PivotTable’s Row and Column Areas to the PivotChart’s Category Axis and Series Axis. The PivotTable toolbar, PivotTable Field List, and PivotTable buttons all work the same on PivotCharts. However, in a PivotChart, you cannot drill down a dimension hierarchy by double-clicking. A double-click on a PivotChart opens a Format dialog box instead. You must use the PivotTable toolbar or drop-down arrow on PivotTable buttons to expand and collapse dimension hierarchy levels.

Excel displays the Chart toolbar with the PivotChart. Use the Chart Type button on the toolbar to choose between Excel’s various chart types for your PivotChart. Note that not all of Excel’s standard chart types are supported with OLAP PivotCharts; for example, you cannot select Scatter or Bubble charts.

NOTE
When you first run the PivotTable and PivotChart Wizard, you can choose to create a PivotChart report. Excel will automatically build a PivotChart and PivotTable. A PivotChart must always have an underlying PivotTable. You can create a PivotTable without a PivotChart, but you can't have a PivotChart without a PivotTable.

TIP FROM
By default, Excel places the PivotChart on a separate chart sheet in your workbook. To place the PivotChart on the same worksheet with the PivotTable, right-click the PivotChart and select Location from the pop-up menu to open the Chart Location dialog box. Click the As Object In radio button and select the worksheet that contains the PivotTable from the drop-down list. Click OK.

SAVING OFFLINE CUBES FROM SERVER CUBES

When connected to a server-based cube with Excel, you can save subsets of the cube’s data to local cubes on your hard disk. You can reload data from the local cubes into Excel PivotTables without connecting to the server. This pack-and-go option enables you to connect your laptop computer to the server while you are in the office, download OLAP cube data, and perform OLAP analysis sessions while on the road or include OLAP analysis as part of offsite presentations. You also can save and send local cube files to other Excel users who don’t have access to your OLAP server. Follow these steps to create an offline cube from a server cube:

1. Create an OLAP PivotTable as you normally would, using an OLAP server as your data source.
2. In the PivotTable toolbar, open the PivotTable drop-down menu and choose Offline OLAP to open the Offline OLAP Settings dialog box. (This option won’t be available unless you have one or more fields displayed in your OLAP PivotTable.)
3. Click the Create Offline Data File button to display the Create Cube File Wizard. The first wizard screen is an information-only screen. Click Next to proceed.
4. Step 2 of the wizard displays a tree of the server cube’s dimensions (see Figure 30.14). Check off the dimension levels you want to include in your offline cube. Click the plus and minus signs to expand and collapse dimension hierarchies to select different detail levels in the dimension, and then click Next.

![Figure 30.14](image)

The server cube might contain too much data for an efficient offline cube. Just check off the parts of the server cube you really need.

5. Step 3 enables you to select the measure fields and dimension data values from the dimension levels you chose in the previous step (see Figure 30.15). Click the plus and minus signs to expand and collapse the tree view and check off the items you want to include in the offline cube. Be sure to select at least one item from each dimension level. Click Next.

![Figure 30.15](image)

You can limit offline cube size by choosing certain measures and only some data values from the dimension levels you previously selected.

6. The last step is to tell Excel where to save the offline cube file. Type the filename or click the Browse button to open the traditional Windows Save As dialog box.

7. Click Finish to begin saving the cube file. Excel displays the Create Cube File - Progress dialog box while it is copying the data. You can click the Stop button to abort the process if it's taking too long.

When Excel is finished creating the local cube file, you are returned to the Offline OLAP Settings dialog box (see Figure 30.16). The Offline OLAP radio button is selected. When you click OK, your PivotTable will be disconnected from the server and working with data from the offline cube file.
Performing OLAP Analysis on Database Data

If you want to switch back to using the server cube, go back to the Offline OLAP Settings dialog box, select the On-Line OLAP radio button, and click OK. You will reconnect to the server cube from which the offline cube was created. You can also return to the Offline OLAP Settings dialog box and choose the Edit Offline Data File button to go back through the Create Cube File Wizard to change the dimensions or measures you included in the offline cube.

Performing OLAP Analysis on Database Data

You can use Excel PivotTables and PivotCharts to analyze relational database data that has not been loaded into cubes on an OLAP server. For example, you might have your own Microsoft Access database that you want to analyze, or perhaps you need to work with data in an Oracle database that wasn’t defined with the requirements when the database administrator built server-based cubes. If you have been using regular Excel PivotTables with a relational data source, you might find that you get better performance if you first build a cube from the data and then use OLAP PivotTables.

Starting the OLAP Cube Wizard

Choose Data, Import External Data, New Database Query, and create a new query or edit one that’s already saved. Be sure the check box to Use the Query Wizard to Create/Edit Queries is selected in the Choose Data Source Dialog. See Chapter 29, “Using Excel with Access and Other Databases,” for details on using the Query Wizard. In the Choose Columns step of the Query Wizard, be sure you include all the database columns from your data source that correspond to the measures and dimension levels you want to include in your cube. When you get to the final step of the Query Wizard (see Figure 30.17), choose Create an OLAP Cube from This Query, and click Finish to launch the OLAP Cube Wizard.

Tip from John

If you are making a query specifically to support cube building, you will get the best performance if you omit any database columns that will not be part of the cube. Also, consider adding a filter to your query so it returns only the data you are interested in analyzing in PivotTables.
The OLAP Cube Wizard starts with a welcome message. After reading this message, click Next to proceed to the first step of the wizard. You can disable the welcome screen for future use by selecting Don’t Show This Screen Again before you click Next.

**SELECTING CUBE MEASURES**

Step 1 of the OLAP Cube Wizard (see Figure 30.18) lets you select the source fields that will be summarized data fields (measures) in your cube. By default, the numeric columns in your query are preselected with the Sum aggregate operator. Deselect any numeric fields that are not truly measures. For example, a store number column might be stored as a numeric data type, and it would be nonsense to sum up the values of the store number in each database record.

You also can decide here whether you want to aggregate your measures with the default Sum operator or if you would rather see the minimum value, maximum value, or count of values when details are rolled up to the next level in dimension hierarchies. For example, if
the sales value of quarters 1, 2, 3, and 4 are $100, $200, $300, and $100, respectively, the sum at the year level will be $700, the minimum for the year would be $100, and the maximum for the year would be $300. To change the summarization type, click the word *Sum* in the Summarize By column and choose Count, Min, or Max from the drop-down list.

You can turn nonnumeric fields into measures by using the Count summarization. For example, counting customer IDs creates a measure field that represents the number of customers. Click the check box to the left of a nonnumeric source field and Excel automatically sets the Summarize By type to Count. Choose Next when you have finished selecting and setting options for all of your cube’s measures.

**Organizing Data into Hierarchies**

Step 2 of the wizard requires you to arrange the remaining source data fields into one or more dimension hierarchies, as shown in Figure 30.19. Group fields logically according to the type of information they contain. For example, a set of fields describing a customer’s address should be grouped into a dimension called customer location, or a set of fields describing various product data should be grouped into a product dimension.

Drag and drop each field from the Source Fields window into the Dimensions window. If you drop the field onto the (Drop Fields Here to Create a Dimension) item, it becomes a new dimension. If you drop it onto an existing dimension, it becomes a level in the dimension hierarchy. Be sure you group your dimension hierarchy levels in the proper order from summary level on top to lowest detail level on bottom. For example, in the Customer Location dimension, Country is the highest summary level, with State Province coming below and City on the bottom.

When you’re finished creating dimensions, click Next.
Congratulations! You've done pretty much what a database administrator does when setting up an OLAP server database for the first time.

Step 3 of the wizard just asks how you would like to save the cube (see Figure 30.20). You can choose to rebuild the cube each time you open a worksheet that references it, or you can build it once and then rebuild it as needed (which is important to do when the source database data is updated).

Figure 30.20
After you've built your new cube database, save it with its data to your local hard disk.

In most cases, you'll want to choose Save a Cube File Containing All Data for the Cube, because this maximizes the speed benefits of using an OLAP cube in the first place.

Type a name for the cube data file (it has a .cub extension), and then click Finish. You’ll be asked to save the OLAP query definition (.oqy extension). Give it a name and click Save. You’ll need to use this query definition file when you want to use the cube.

Next, you’ll see the OLAP Cube Wizard running the query and creating the cube. A status box appears, saying that Excel is creating an offline cube.

NOTE
It’s normal for this process to take a few minutes, particularly if you’re summarizing a large query.
After the local cube has been created, Excel pops you into Step 3 of the PivotTable and PivotChart Wizard so you can immediately create an OLAP PivotTable based on your new cube (refer to Figure 30.9).

**Using the Cube**

If you chose not to create an OLAP PivotTable as part of the process of creating your local OLAP cube, or if you want to reuse the same cube in another PivotTable, you'll need to tell Excel to use the cube for OLAP data. If you think back to when you first created an OLAP PivotTable, you'll probably remember that the PivotTable and PivotChart Wizard gives you the option of using a local OLAP cube instead of an OLAP server.

To take advantage of this option, follow these steps:

1. Choose Data, PivotTable, and PivotChart Report to start the PivotTable and PivotChart Wizard.
2. In the Step 1 of 3 dialog box, choose External Data Source and click Next.
3. In the Step 2 of 3 dialog box, click the Get Data button.
4. When the Choose Data Source dialog box opens, click the OLAP Cubes tab. There you'll see listed the OLAP query definition you created when you built the cube in the OLAP Cube Wizard. Select it and click OK. You'll then be back in the PivotTable and PivotChart Wizard, Step 2 of 3.
5. The rest of this should be old hat to you. Click Next; then decide where you want the OLAP PivotTable to go and click Finish.

**Excel in Practice**

OLAP PivotTables can display an enormous amount of detail—sometimes too much. For example, it's hard to see in the promotions analysis in Figure 30.21 what data is really important and what's just background noise. One of the goals of OLAP is to help you quickly find the hot spots in your data. You can apply Excel features to OLAP PivotTables to zero in on the important information and reveal hidden treasure in your data.

To turn the raw data into useful information, you can use several techniques:

- Discover the top performers by placing the cell pointer in the Promotion Name column and choosing PivotTable, Field Settings, Advanced from the PivotTable toolbar (or right-clicking anywhere in the column and choosing Field Settings, Advanced) to open the PivotTable Field Advanced Options dialog box (see Figure 30.22). Choose Descending under AutoSort Options and select Store Sales from the Using Field drop-down list to sort promotions by descending store sales.
Cut down on the number of promotions by again opening the PivotTable Field Advanced Options dialog box and choosing On under Top 10 AutoShow to limit the list to the top 10 promotions. Now you know which promotions were the top 10 revenue earners.

Place the cell pointer in a Store Sales data cell, and then choose PivotTable, Field Settings, Options, Show Data As. Set Store Sales to display as % of row, instead of an absolute value.

Finally, select the entire PivotTable (PivotTable, Select, Entire Table) and use the Format, Conditional Formatting command to mark values between 50% and 99% in red font with shaded background. Now, out of the top promotions, you instantly can tell where each worked the best.
For example, it's now easy to see in Figure 30.23 that You Save Days is a promotion to repeat in California but not worth running again in Oregon or Washington. However, Washington customers responded extremely well to the Super Savers promotion, making it a winner in that state. Note that the formatting changes made previously will persist as you continue to slice and dice through the data in your OLAP PivotTable. Excel will automatically reapply the formatting whenever the PivotTable data is refreshed.