



MASSTRANSIT DATABASE ANALYSIS

Using Microsoft Excel And ODBC To Analyze
MySQL Data

AUTHOR: PETE CODY

petecody@grouplogic.com

31 May 2007

Document Revision History:

Date:	Author:	Comment:
19 Apr 2007	PMC	Created
22 May 2007	PMC	Modified screenshots; added chart reporting.
26 May 2007	PMC	Removed cut/paste tags.
29 May 2007	PMC	Formatting Edits

Summary:

This document explains the process in which a MassTransit user may utilize the MySQL ODBC Connector in conjunction with Microsoft Excel and MySQL Server in order to analyze MassTransit data.

Note to Reader: Readers are encouraged to send any feedback, questions, comments, or concerns to support@grouplogic.com. Group Logic would also like to know more about your reporting interests and needs. Please submit these reporting requests to mtsales@grouplogic.com.

Contents:

Installation of MyODBC Database Connector:	3
Configuring the MyODBC Database Connector:	3
Connecting Microsoft Excel to MySQL:	5
Creating A 'Top Senders' Report Within Excel:	8
Creating A 'Top Receivers' Report Within Excel:	12
Creating A 'Daily Trend – Total FILES Transferred By Date' Report Within Excel:.....	16
Creating A 'Daily Trend – Total BYTES Transferred By Date' Report Within Excel:...	19

Installation of MyODBC Database Connector:

Open DataBase Connectivity (ODBC) is a standardized interface which may be used to access a database from an application. MySQL, the database backbone of MassTransit, provides a 'connector' which enables various applications such as Microsoft Excel to analyze specific datasets. In this section, we will install the MyODBC Database Connector.

Download the MySQL Connector (version ODBC 3.5.1):
<http://dev.mysql.com/downloads/connector/odbc/3.51.html>

For convenience, I have created a **TinyURL** path that links directly to the MySQL Connector version 3.5.1 MSI Installer (Windows): <http://tinyurl.com/348y8k>

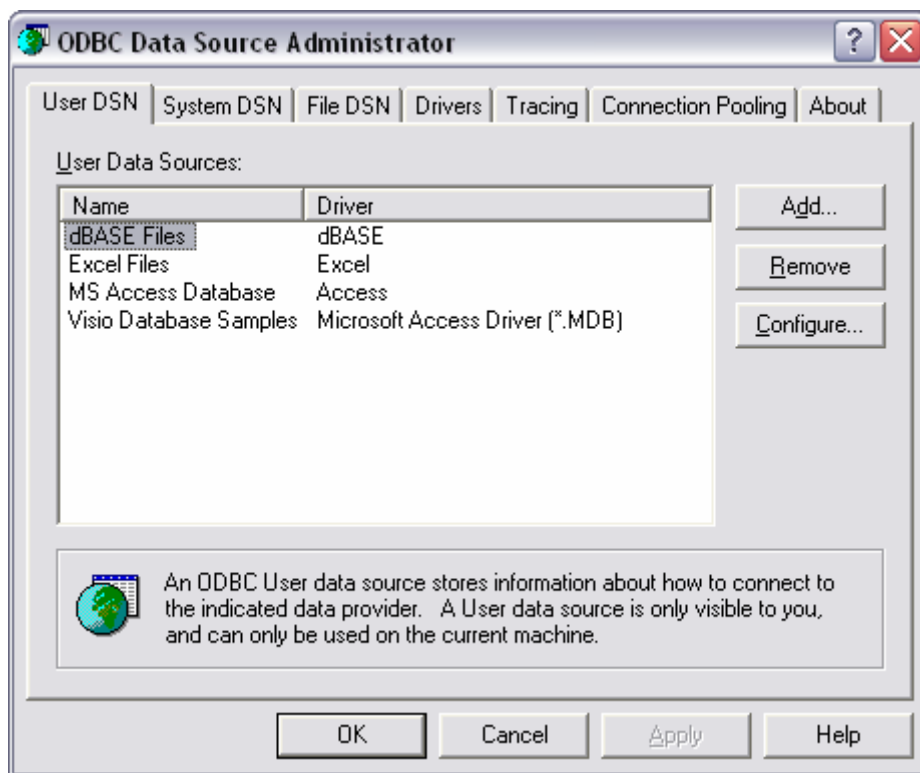
After the download has completed, click on the installer icon to begin the installation process. When prompted, select the **Typical** setup option. Click the **Install** button to begin the installation process, and then click the **Finish** button when the installation completes.

Configuring the MyODBC Database Connector:

In this section, we will configure the MyODBC Connector

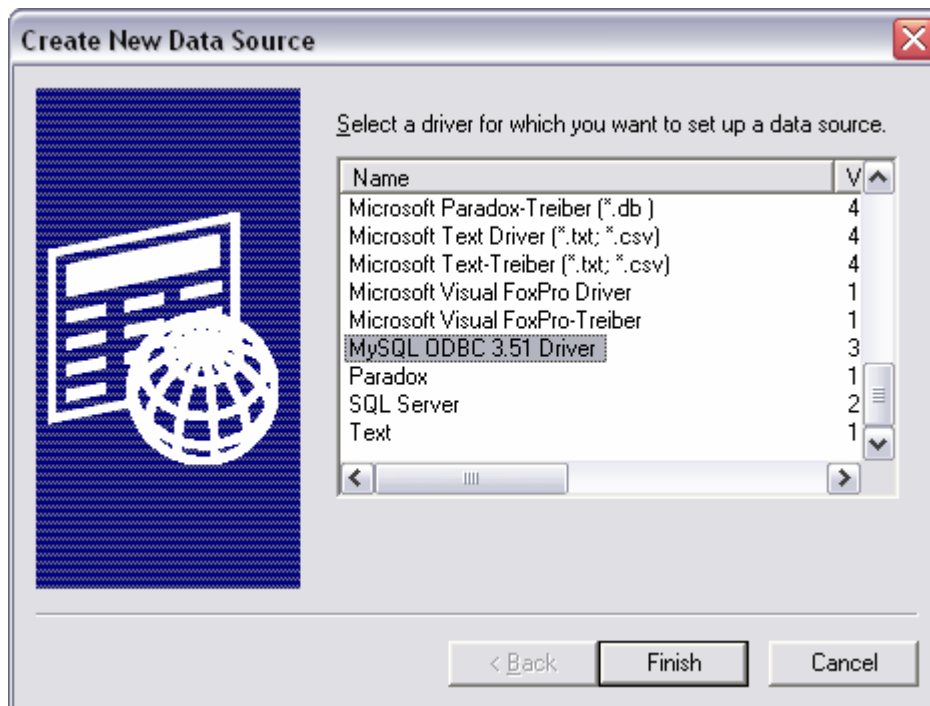
Open the ODBC Data Source Administrator:

Settings > Control Panel > Administrative Tools > ODBC Data Source Administrator



The ODBC Data Source Administrator

Select the "User DSN" tab, and click on the **Add** button to create a new data source. Double-click the item "MySQL ODBC driver" to configure the MySQL ODBC driver.



Create New Data Source within the ODBC Administrator

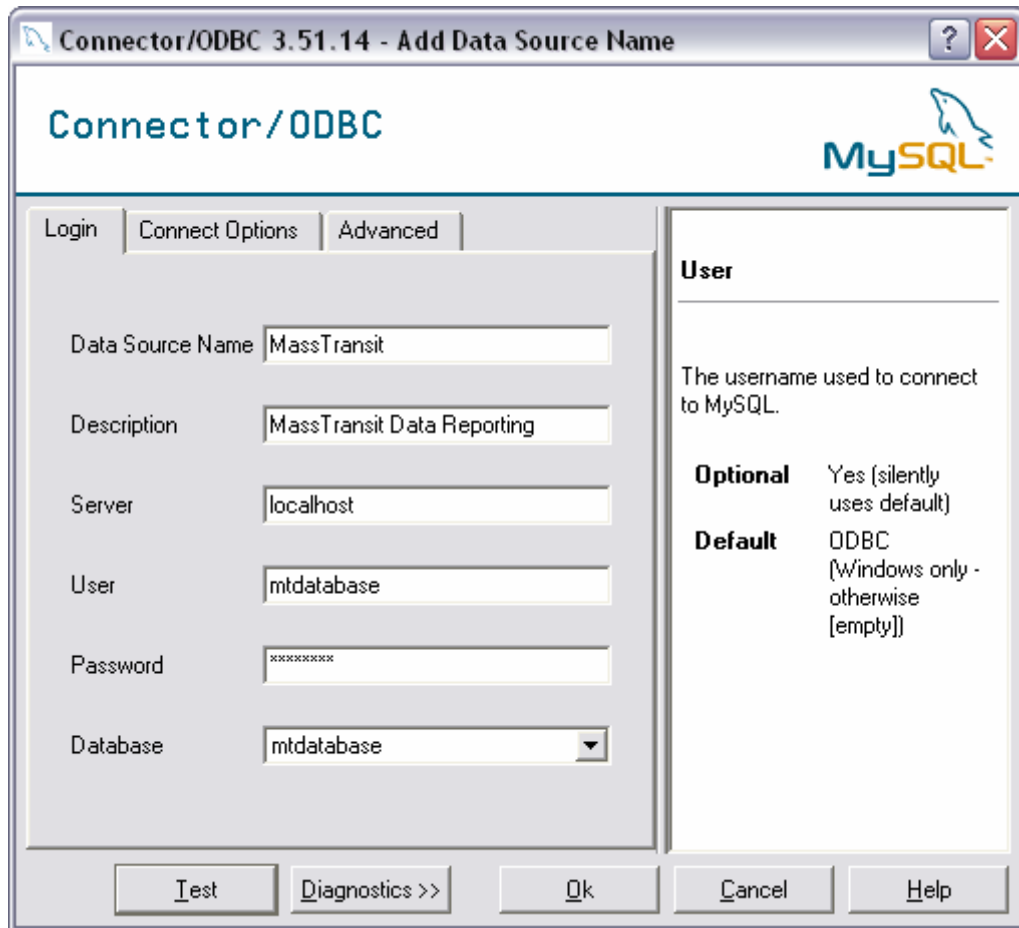
Within the **Data Source Name** field, enter a unique name that will be used to identify the connection to the MassTransit database within MySQL.

Next, within the **Description** field, enter a brief sentence that describes the purpose of this ODBC connection.

Next, enter the Fully Qualified Domain Name (FQDN) or IP address of the MySQL server that houses the MassTransit databases within the **Server** field. If this connection is being installed on the MassTransit server itself, enter "localhost" without quotes.

Next, provide a valid username and password in the **User** and **Password** fields. This must be an existing user that has been configured within the MySQL Administrator. This user must have permissions to view the contents of the MTDatabase schema.

Finally, select MassTransit's database in the **Database** drop-down box. If your connection to the MySQL server is properly configured, the drop-down box should automatically populate with all databases that are visible to that user. (Note: In a default configuration, the MassTransit Database will be named "mtdatabase." You may review your *MassTransit Engine.cfg* file to locate your MassTransit database name.)



Configuring a MySQL Data Source Within the ODBC Administrator

After these options have been configured, press the **Test** button to confirm a successful connection to the MySQL Server can be made. If a connection cannot be made, double check the configuration settings above.

Press the **OK** button to save the changes, and return to the main ODBC Data Source Administrator window. At this point, your MassTransit connection should be listed among any other previously configured data sources.

Connecting Microsoft Excel to MySQL:

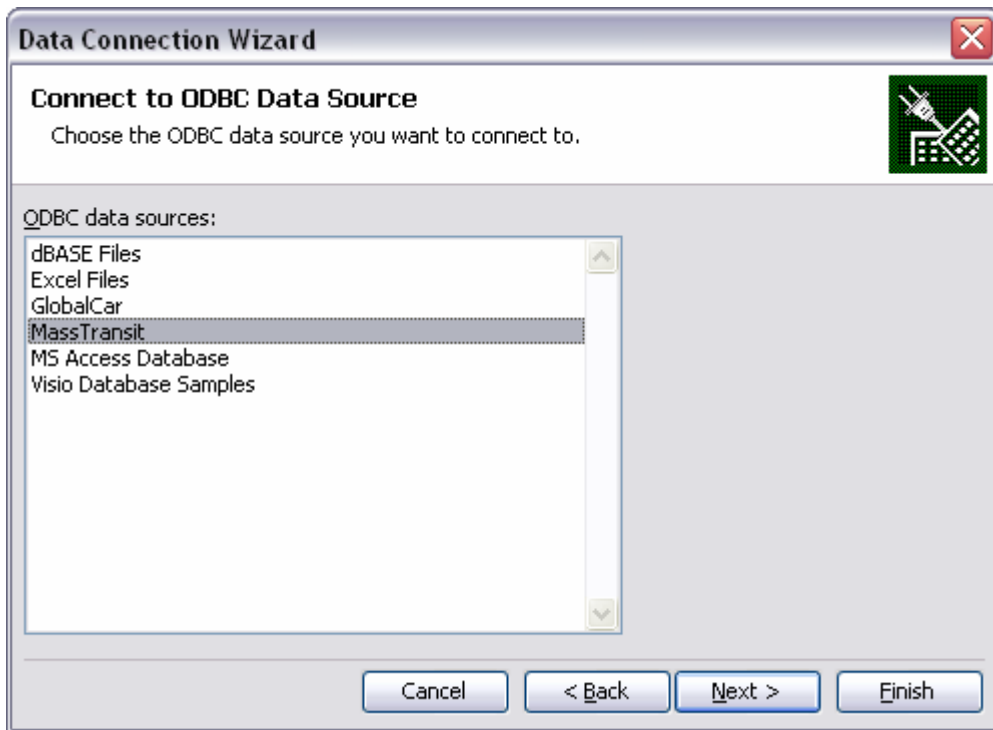
To access the MassTransit data from within Microsoft Excel, please complete the process below.

Open the Excel application and navigate to **Data > Import External Data > Import Data**.

Next, select the option **+Connect to New Data Source.odc**

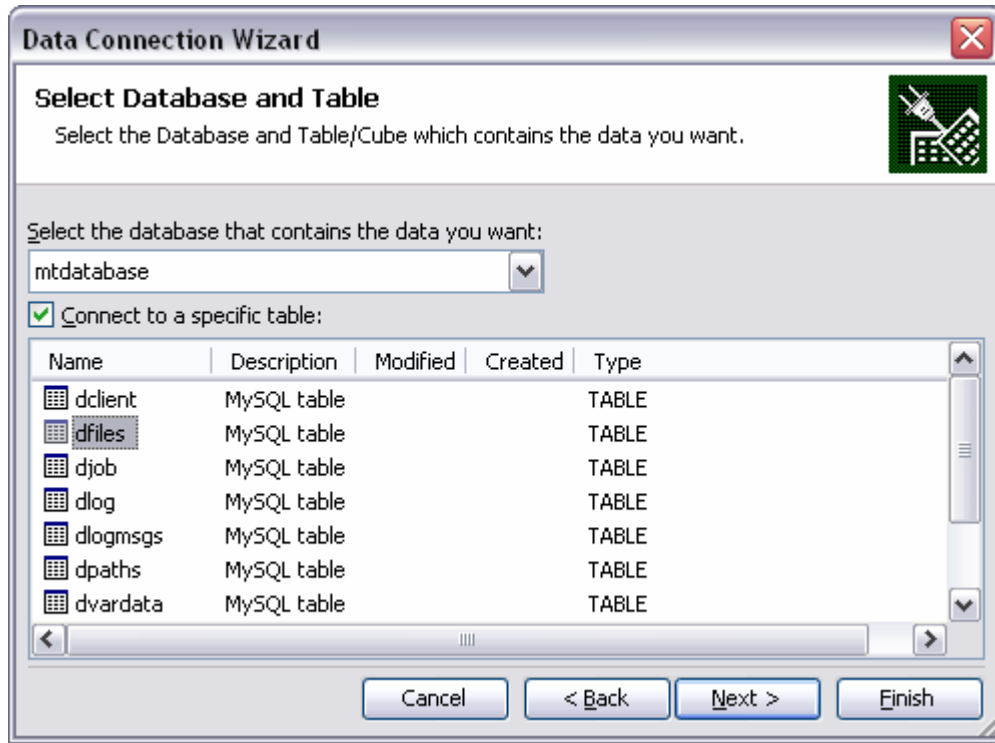
When the “Welcome to the Data Connection Wizard” window appears, select the **ODBC DSN** option, and click the **Next** button to connect to the ODBC data source that was created in the previous section.

Next, from the window titled “Connect to ODBC Data Source,” choose the DSN that was created in the previous section. In this case, it is “MassTransit.” Press the **Next Button** to begin selecting the databases and tables to begin reporting.



Connecting to the MassTransit Data Source

Now that you have connected to the MassTransit Data Source, you will be presented with a dialog that allows you to select the various MySQL tables.



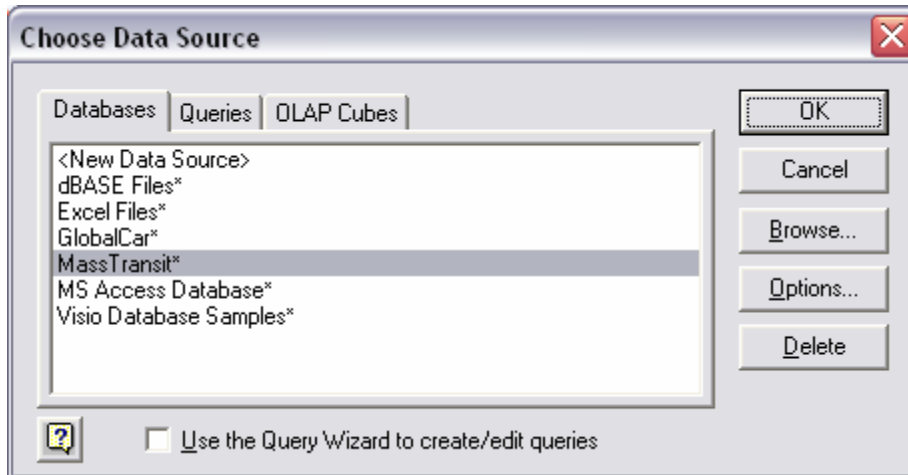
MassTransit Tables Presented Within Excel's Data Connection Wizard

Select a table for reporting, and then click the **Next** button. A window titled "Save Data Connection File and Finish" should appear. Press the **Finish** button to save the MassTransit data connection.

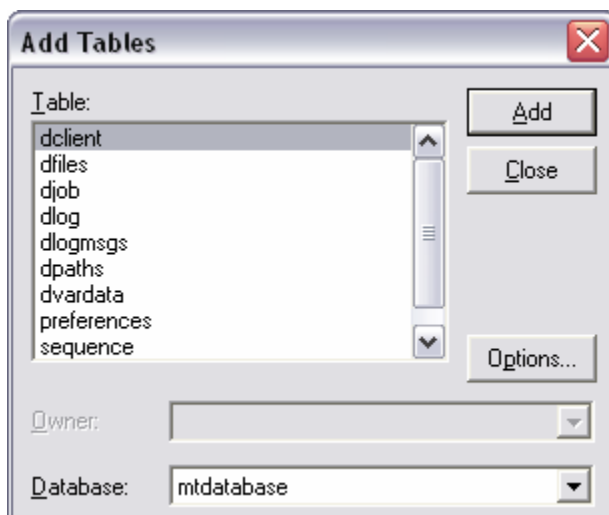
Creating A 'Top Senders' Report Within Excel:

In this section, we will create a report that shows the MassTransit contacts that have **sent** the most number of files.

1. Open Microsoft Excel, and create a new blank workbook (File > New > Blank Workbook.)
2. From the Data pull down menu, select Import External Data > New Database Query.
3. The 'Choose Data Source' window will appear.
4. Select "MassTransit*" and uncheck "Use the Query Wizard to create/edit queries" and click OK.



5. After you have selected "OK," the "Microsoft Query" window will appear. The "Add Tables" window should be in the forefront. NOTE: If the "Add Tables" window does not automatically appear, you may open it by selecting "Table" > "Add Tables..."
6. The Top Sender data that we will be requesting is found in the **dclient** table.
7. Within the "Add Tables" window, double click on the **dclient** table.



8. Now, click on the red "X" in the upper-right corner to close the "Add Tables" window.

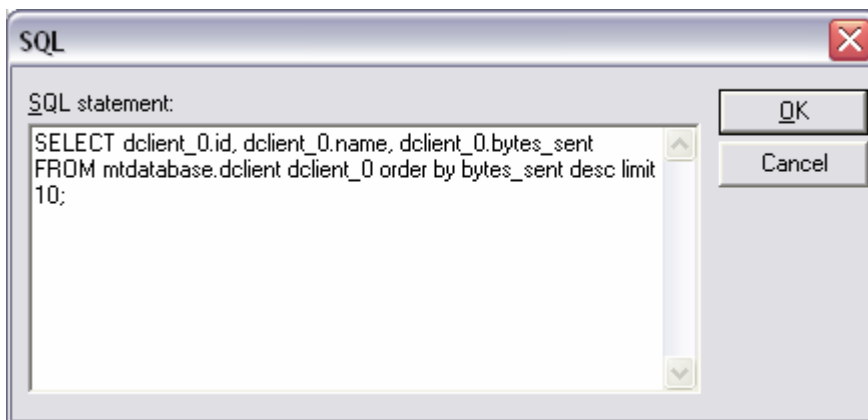
Now that the **dclient** table has been selected, we can now build our SQL query statement to select the top ten senders.

- Click on the "SQL" button found towards the upper-right hand corner of the Microsoft Query window. This will open the SQL statement window. (Alternatively, this window can be opened by selecting the "View" pulldown menu, and then selecting "SQL..."
- Within the SQL window, please enter the following SQL statement::

SQL statement to retrieve Top Sender Data from the MassTrasnsit Database:

```
SELECT dclient_0.id, dclient_0.name, dclient_0.bytes_sent  
FROM mtdatabase.dclient dclient_0 order by bytes_sent desc limit 10;
```

Note: The 'order by bytes_sent desc' statement will sort the displayed data by total bytes sent, in descending order. The 'limit 10' parameter *shown in Red* limits the returned results to show the top ten senders, based on number of files that have been sent. That value can be increased or decreased to limit the scope of the report.

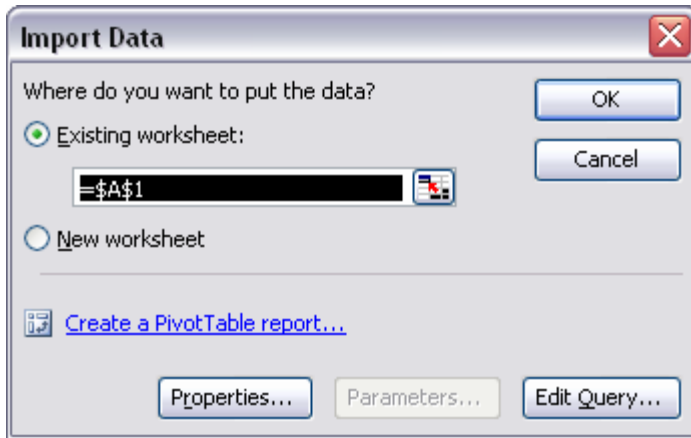


The Top Sender query as it should appear within the SQL query window

- After the SQL statement has been entered into the query window, click the OK button to execute the query. The returned results will be displayed in a spreadsheet-like format.
- Once the returned data has been displayed, import it to your workbook by selecting the "File" pulldown menu, and then selecting "Return Data to Microsoft Office Excel."
- The "Import Data" window will appear. Select the cell where the imported data will be imported. Note that data will expand vertically and horizontally from this point, dependent on the number of rows and columns in your exported result set.
- Click "OK" to import the data.

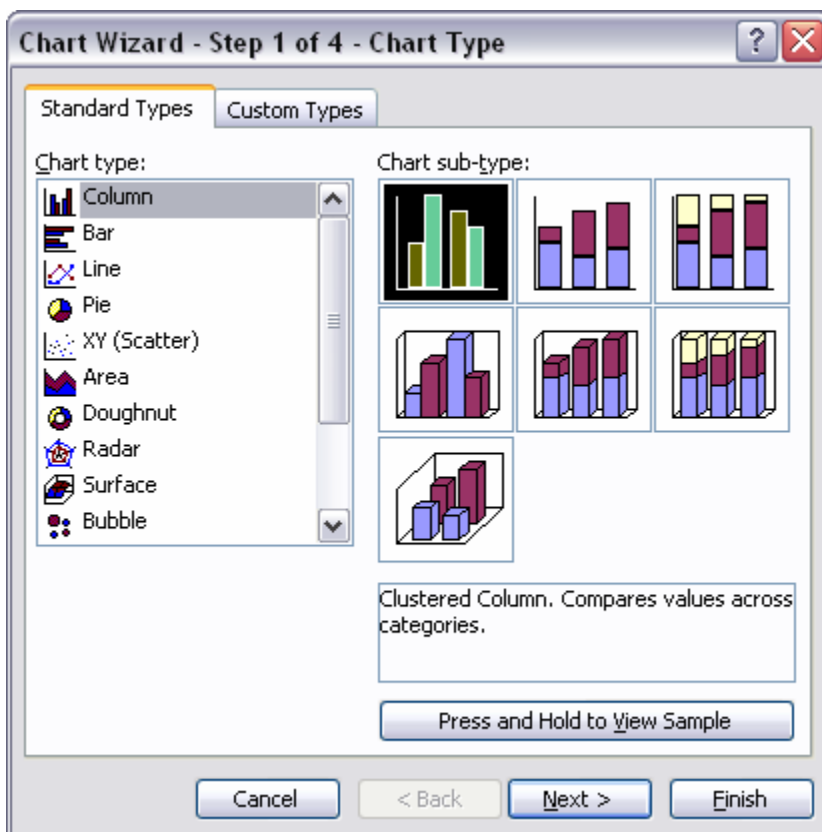
After the import step has been completed, you will be returned to the blank worksheet from where you started. An "Import Data" dialog box will be presented, allowing one to select where the data will be placed within the workbook.

- Select the row and column where the data will begin to be imported. This location represents the upper-left hand corner of the data to be displayed.
- Click the "OK" button to paste data into the worksheet.



The Microsoft Excel Import Data Dialog

17. The data will now be placed into the spreadsheet in the location which you have specified. The first row of imported data will contain the field names of the data imported.
18. Next, from the "Insert" pull-down menu, select the "Chart" option.
19. When the "Chart Wizard" appears, select the type of chart to best graphically represent your data. (in most cases, Column chart type is sufficient.)



The Microsoft Excel Chart Wizard – Select Column Chart type, and Clustered Column sub-type

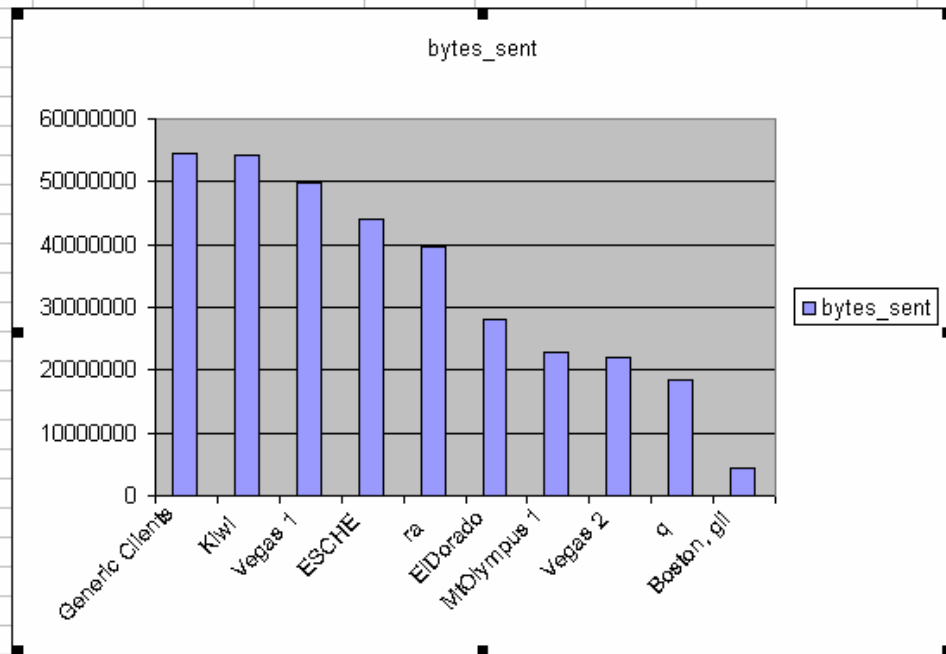
20. After selecting the appropriate Chart type, click the Next button to continue.
21. The Chart Wizard will then prompt to enter the Data range. This is the data that will be included in your graphical chart. Using your cursor, highlight the area that you wish to include in your report.

	A	B	C	D
1	id	name	bytes_sent	
2	1	Generic Clients	54623423	
3	52	Kiwi	54355211	
4	53	Vegas 1	49850021	
5	54	ESCHE	43903882	
6	61	ra	39724041	
7	58	EIDorado	28014321	
8	65	MtOlympus 1	22717405	
9	56	Vegas 2	21908432	
10	60	q	18324524	
11	66	Boston_gli	4325222	
12				

Microsoft Excel Chart Wizard – Selecting Chart Data Area

22. The Source Data window will appear, showing a mini-representation of the data you have selected. **Be sure to include the column headers, as these will represent the chart titles and data labels.** Click the “Next” button to proceed.
23. Click the “Finish” button to paste your chart into your worksheet.

[014321](#)
[717405](#)
[908432](#)
[324524](#)
[325222](#)

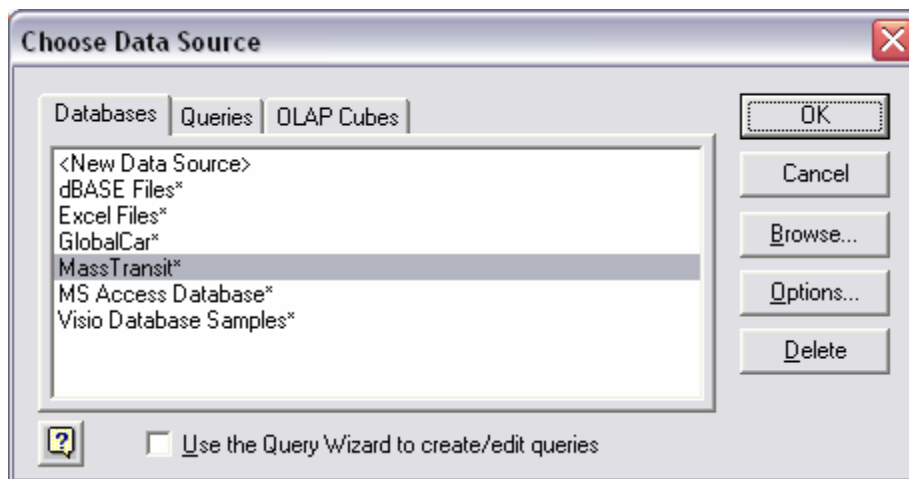


Example of a Completed "Bytes Sent" report within an Excel Spreadsheet

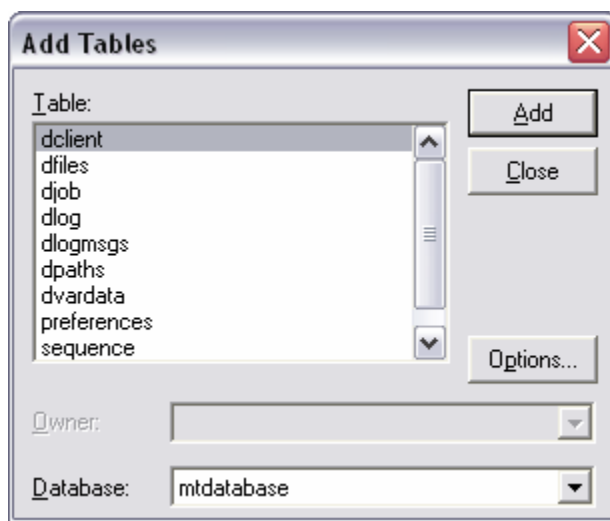
Creating A 'Top Receivers' Report Within Excel:

In this section, we will create a report that shows the MassTransit contacts that have **received** the most number of files.

1. Open Microsoft Excel, and create a new blank workbook (File > New > Blank Workbook.)
2. From the Data pull down menu, select Import External Data > New Database Query.
3. The 'Choose Data Source' window will appear.
4. Select "MassTransit*" and uncheck "Use the Query Wizard to create/edit queries" and click OK.



5. After you have selected "OK," the "Microsoft Query" window will appear. The "Add Tables" window should be in the forefront. NOTE: If the "Add Tables" window does not automatically appear, you may open it by selecting "Table" > "Add Tables..."
6. The Top Receiver data that we will be requesting is found in the **dclient** table.
7. Within the "Add Tables" window, double click on the **dclient** table.



8. Now, click on the red "X" in the upper-right corner to close the "Add Tables" window.

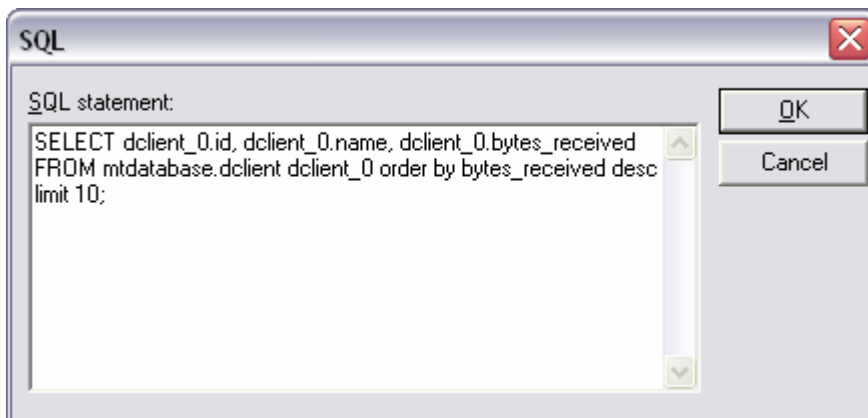
Now that the **dclient** table has been selected, we can now build our SQL query statement to select the top ten receivers.

9. Click on the "SQL" button found towards the upper-right hand corner of the Microsoft Query window. This will open the SQL statement window. (Alternatively, this window can be opened by selecting the "View" pull down menu, and then selecting "SQL..."
10. Within the SQL window, please enter the following SQL statement::

SQL statement to retrieve Top Receiver Data from the MassTransit Database:

```
SELECT dclient_0.id, dclient_0.name, dclient_0.bytes_received  
FROM mtdatabase.dclient dclient_0 order by bytes_received desc limit 10;
```

Note: The 'order by bytes_received desc' statement will sort the displayed data by total bytes received, in descending order. The 'limit 10' parameter *shown in Red* limits the returned results to show the top ten recipients, based on number of files that have been received. That value can be increased or decreased to limit the scope of the report.

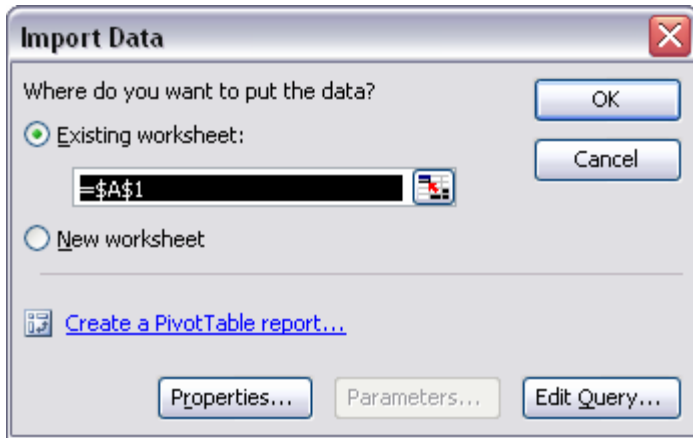


The Top Receiver query as it should appear within the SQL query window

11. After the SQL statement has been entered into the query window, click the OK button to execute the query. The returned results will be displayed in a spreadsheet-like format.
12. Once the returned data has been displayed, import it to your workbook by selecting the "File" pull down menu, and then selecting "Return Data to Microsoft Office Excel."
13. The "Import Data" window will appear. Select the cell where the imported data will be imported. Note that data will expand vertically and horizontally from this point, dependent on the number of rows and columns in your exported result set.
14. Click "OK" to import the data.

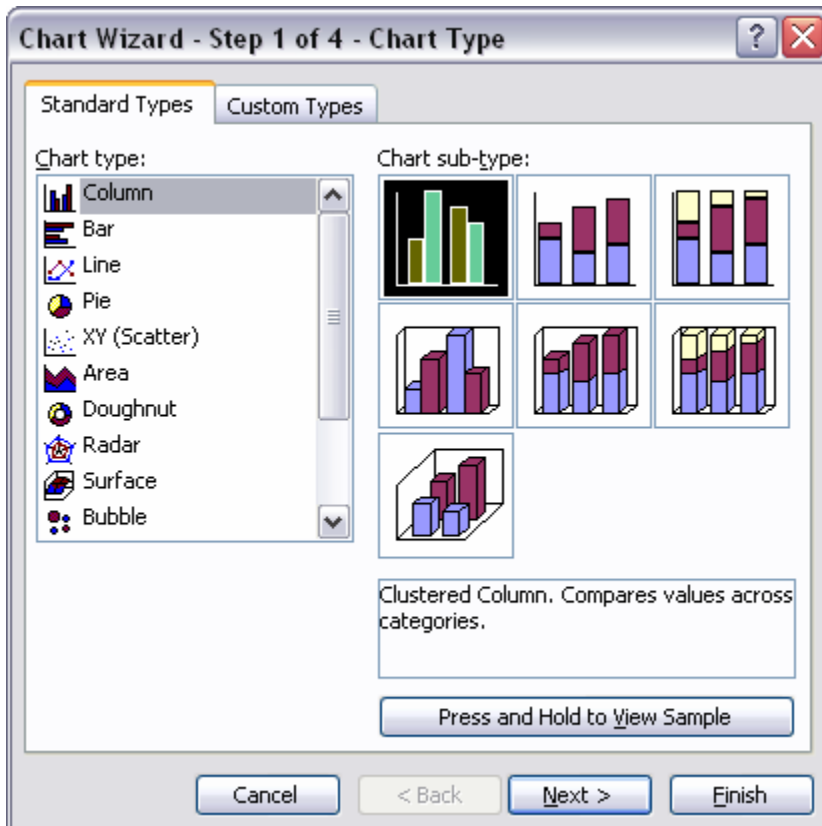
After the import step has been completed, you will be returned to the blank worksheet from where you started. An "Import Data" dialog box will be presented, allowing one to select where the data will be placed within the workbook.

15. Select the row and column where the data will begin to be imported. This location represents the upper-left hand corner of the data to be displayed.
16. Click the "OK" button to paste data into the worksheet.



The Microsoft Excel Import Data Dialog

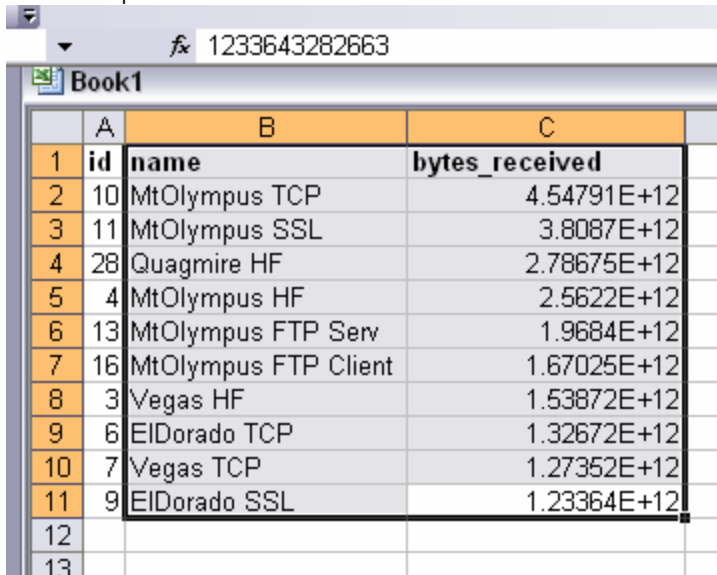
17. The data will now be placed into the spreadsheet in the location which you have specified. The first row of imported data will contain the field names of the data imported.
18. Next, from the "Insert" pull-down menu, select the "Chart" option.
19. When the "Chart Wizard" appears, select the type of chart to best graphically represent your data. (in most cases, Column chart type is sufficient.)



The Microsoft Excel Chart Wizard – Select Column Chart type, and Clustered Column sub-type

20. After selecting the appropriate Chart type, click the Next button to continue.

21. The Chart Wizard will then prompt to enter the Data range. This is the data that will be included in your graphical chart. Using your cursor, highlight the area that you wish to include in your report.

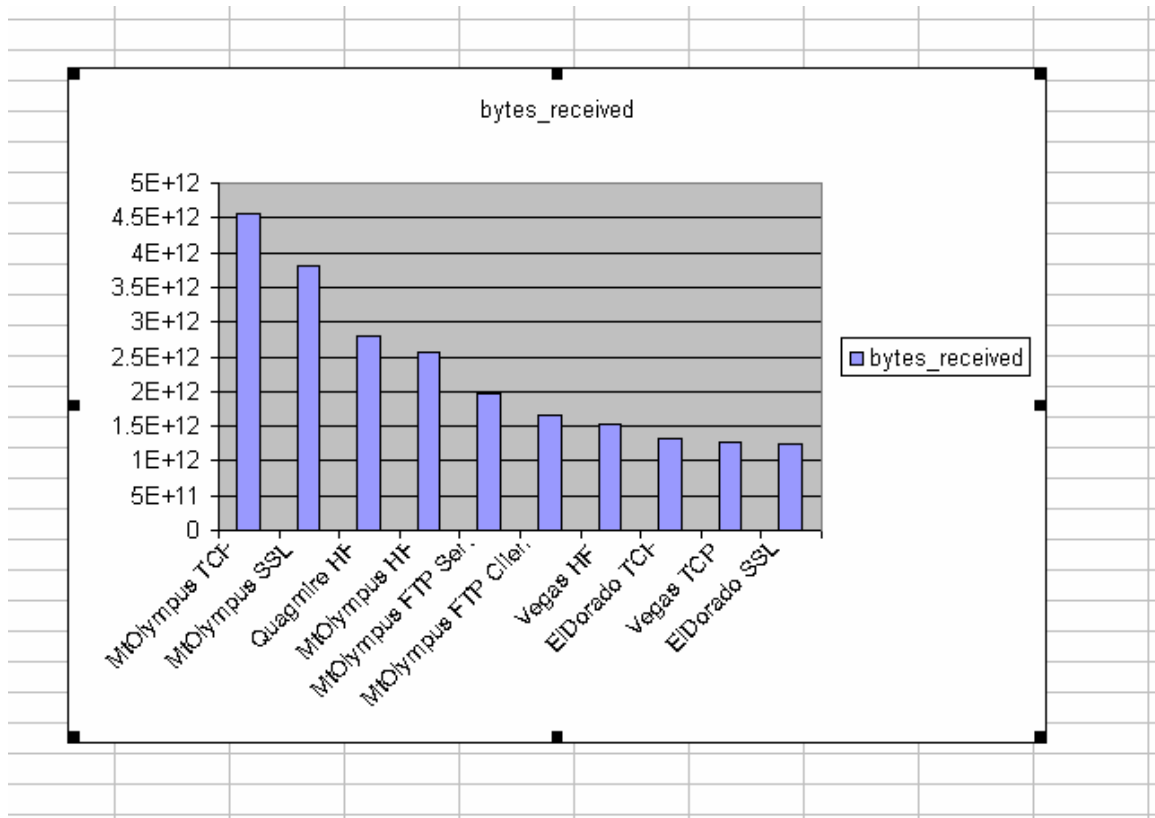


The screenshot shows a Microsoft Excel spreadsheet with a formula bar at the top containing the value 1233643282663. The spreadsheet is titled 'Book1' and contains a table with three columns: 'id', 'name', and 'bytes_received'. The data is as follows:

	A	B	C
1	id	name	bytes_received
2	10	MtOlympus TCP	4.54791E+12
3	11	MtOlympus SSL	3.8087E+12
4	28	Quagmire HF	2.78675E+12
5	4	MtOlympus HF	2.5622E+12
6	13	MtOlympus FTP Serv	1.9684E+12
7	16	MtOlympus FTP Client	1.67025E+12
8	3	Vegas HF	1.53872E+12
9	6	EIDorado TCP	1.32672E+12
10	7	Vegas TCP	1.27352E+12
11	9	EIDorado SSL	1.23364E+12
12			
13			

Microsoft Excel Chart Wizard – Selecting Chart Data Area

22. The Source Data window will appear, showing a mini-representation of the data you have selected. **Be sure to include the column headers, as these will represent the chart titles and data labels.** Click the “Next” button to proceed.
23. Click the “Finish” button to paste your chart into your worksheet.

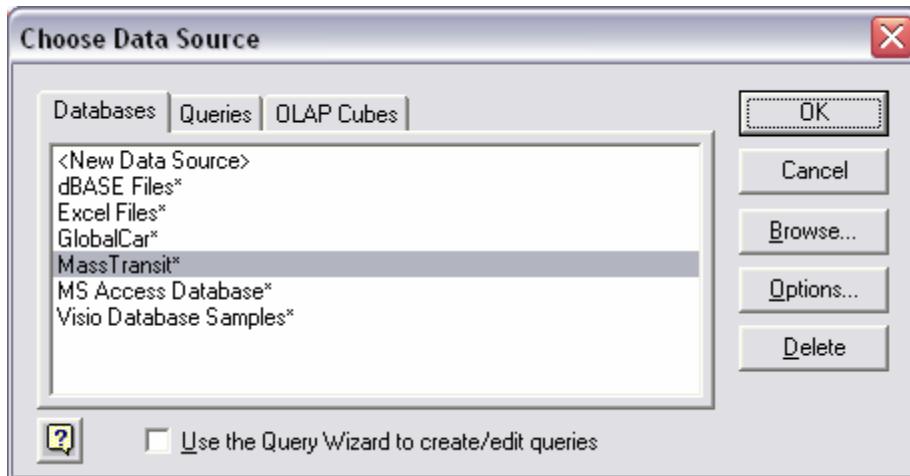


Example of a Completed "Bytes Received" report within an Excel Spreadsheet

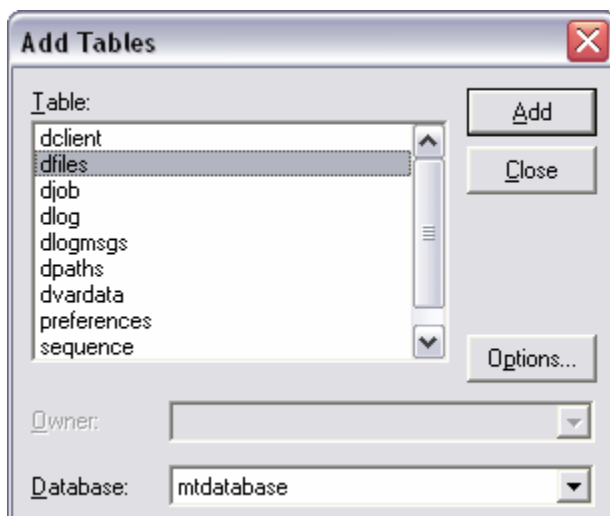
Creating A 'Daily Trend - Total FILES Transferred By Date' Report Within Excel:

In this section, we will create a report that will show the total number of files that the MassTransit server has sent and received on a day-by-day basis.

1. Open Microsoft Excel, and create a new blank workbook (File > New > Blank Workbook.)
2. From the Data pulldown menu, select Import External Data > New Database Query.
3. The 'Choose Data Source' window will appear.
4. Select "MassTransit*" and uncheck "Use the Query Wizard to create/edit queries" and click OK.



5. After you have selected "OK," the "Microsoft Query" window will appear. The "Add Tables" window should be in the forefront. NOTE: If the "Add Tables" window does not automatically appear, you may open it by selecting "Table" > "Add Tables..."
6. The "total transfer" data that we will be requesting is found in the **dfiles** table.
7. Within the "Add Tables" window, double click on the **dfiles** table.



8. Now, click on the red "X" in the upper-right corner to close the "Add Tables" window.

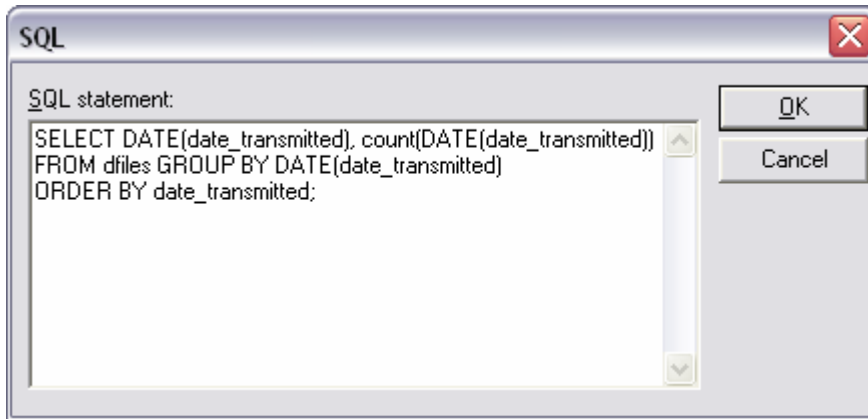
Now that the **dfiles** table has been selected, we can now build our SQL query statement to pull data necessary to create our "total files transferred" report.

9. Click on the "SQL" button found towards the upper-right hand corner of the Microsoft Query window. This will open the SQL statement window. (Alternatively, this window can be opened by selecting the "View" pull down menu, and then selecting "SQL..."
10. Within the SQL window, please enter the following SQL statement::

SQL statement to retrieve Total Files Transferred data:

```
SELECT DATE(date transmitted), count(DATE(date transmitted))
FROM dfiles GROUP BY DATE(date_transmitted)
ORDER BY date transmitted;
```

Note: This statement selects the total number of files send and received by MassTransit, and organizes by date. The DATE functions trim the time stamp to report only the Date, as opposed to Date and Time. The COUNT statement, in combination with the GROUP BY statement will calculate the number of files that have been transmitted (sent and received) on a particular date. The ORDER BY statement shows that data will be displayed in order of date transmitted (ascending order).



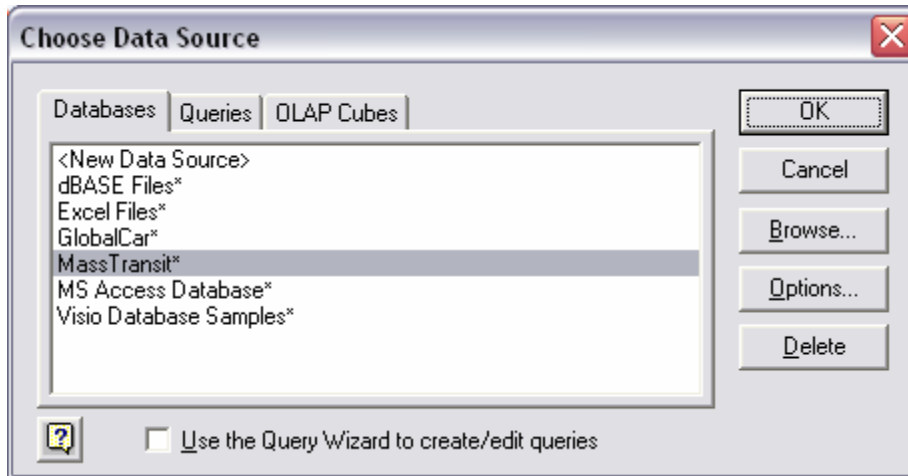
The Total Files Transferred By Date query as it should appear within the SQL query window

11. After the SQL statement has been entered into the query window, click the OK button to execute the query. The returned results will be displayed in a spreadsheet-like format.
12. Once the returned data has been displayed, import it to your workbook by selecting the "File" pulldown menu, and then selecting "Return Data to Microsoft Office Excel."
13. The "Import Data" window will appear. Select the cell where the imported data will be imported. Note that data will expand vertically and horizontally from this point, dependent on the number of rows and columns in your exported result set.
14. Click "OK" to import the data.

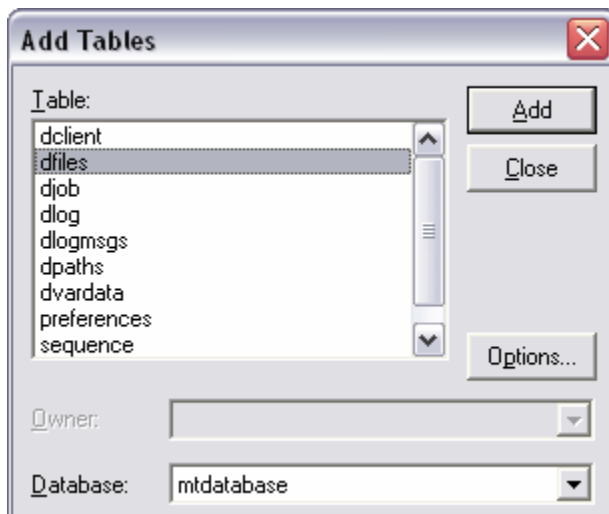
Creating A 'Daily Trend – Total BYTES Transferred By Date' Report Within Excel:

In this section, we will create a report that will show the total number of bytes that the MassTransit server has sent and received on a day-by-day basis.

1. Open Microsoft Excel, and create a new blank workbook (File > New > Blank Workbook.)
2. From the Data pulldown menu, select Import External Data > New Database Query.
3. The 'Choose Data Source' window will appear.
4. Select "MassTransit*" and uncheck "Use the Query Wizard to create/edit queries" and click OK.



5. After you have selected "OK," the "Microsoft Query" window will appear. The "Add Tables" window should be in the forefront. NOTE: If the "Add Tables" window does not automatically appear, you may open it by selecting "Table" > "Add Tables..."
6. The "total bytes transferred" data that we will be requesting is found in the **dfiles** table.
7. Within the "Add Tables" window, double click on the **dfiles** table.



8. Now, click on the red "X" in the upper-right corner to close the "Add Tables" window.

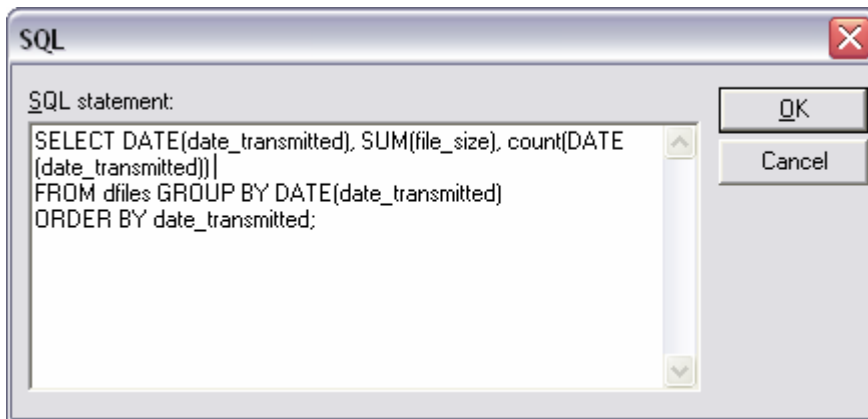
Now that the **dfiles** table has been selected, we can now build our SQL query statement to pull data necessary to create our "total bytes transferred" report.

- Click on the "SQL" button found towards the upper-right hand corner of the Microsoft Query window. This will open the SQL statement window. (Alternatively, this window can be opened by selecting the "View" pulldown menu, and then selecting "SQL..."
- Within the SQL window, please enter the following SQL statement::

SQL statement to retrieve Total BYTES Transferred data:

```
SELECT DATE(date_transmitted), SUM(file_size),  
count(DATE(date_transmitted))  
FROM dfiles GROUP BY DATE(date_transmitted)  
ORDER BY date_transmitted;
```

Note: This statement selects the total number of files send and received by MassTransit, and organizes by date. The DATE functions trim the time stamp to report only the Date, as opposed to Date and Time. The SUM statement will calculate the total number of bytes transferred. The COUNT statement, in combination with the GROUP BY statement will calculate the number of files that have been transmitted (sent and received) on a particular date. The ORDER BY statement shows that data will be displayed in order of date transmitted (ascending order).

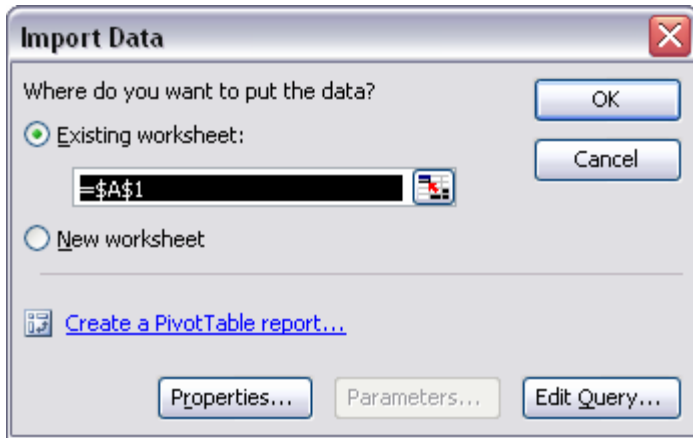


The Total Bytes Transferred By Date query as it should appear within the SQL query window

- After the SQL statement has been entered into the query window, click the OK button to execute the query. The returned results will be displayed in a spreadsheet-like format.
- Once the returned data has been displayed, import it to your workbook by selecting the "File" pulldown menu, and then selecting "Return Data to Microsoft Office Excel."
- The "Import Data" window will appear. Select the cell where the imported data will be imported. Note that data will expand vertically and horizontally from this point, dependent on the number of rows and columns in your exported result set.
- Click "OK" to import the data.

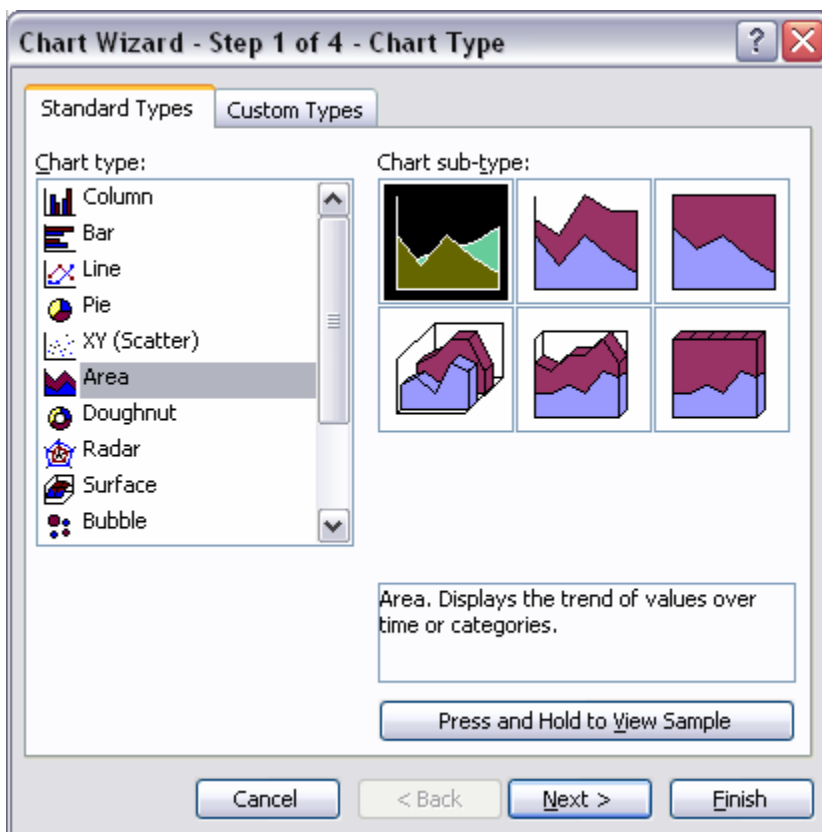
After the import step has been completed, you will be returned to the blank worksheet from where you started. An "Import Data" dialog box will be presented, allowing one to select where the data will be placed within the workbook.

- Select the row and column where the data will begin to be imported. This location represents the upper-left hand corner of the data to be displayed.
- Click the "OK" button to paste data into the worksheet.



The Microsoft Excel Import Data Dialog

17. The data will now be placed into the spreadsheet in the location which you have specified. The first row of imported data will contain the field names of the data imported.
18. Next, from the "Insert" pull-down menu, select the "Chart" option.
19. When the "Chart Wizard" appears, select the type of chart to best graphically represent your data. (in this case, the Area chart type should be appropriate.)



The Microsoft Excel Chart Wizard – Select Area Chart type, and Standard sub-type

20. After selecting the appropriate Chart type, click the Next button to continue.
21. The Chart Wizard will then prompt to enter the Data range. This is the data that will be included in your graphical chart. Using your cursor, highlight the area that you wish to include in your report.

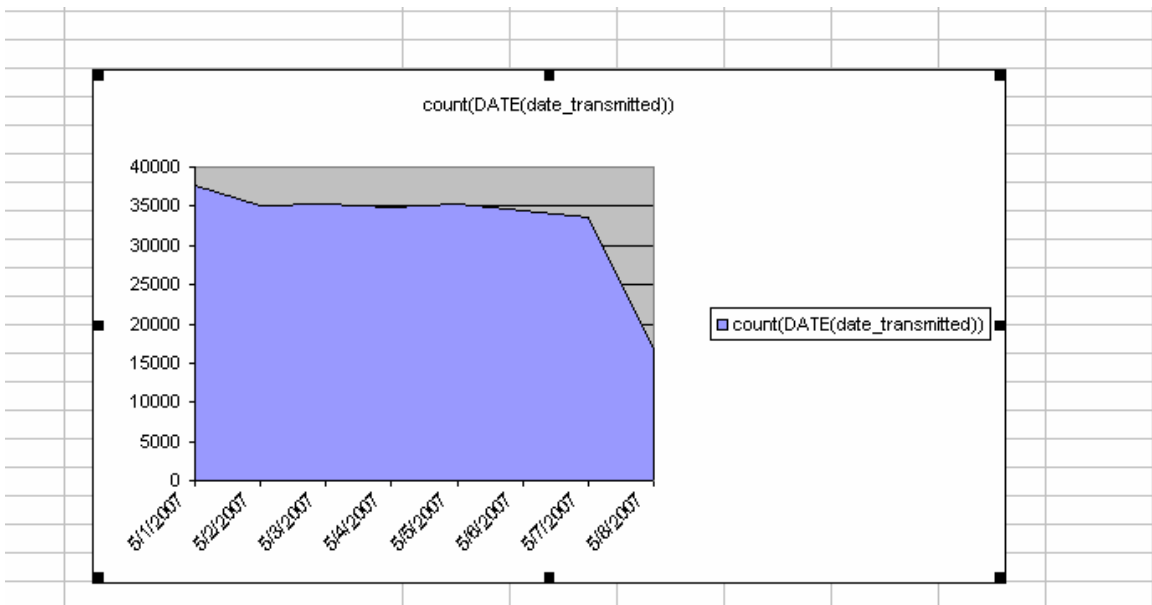
Note: For the "Total Files Transferred By Date" report, you will want to select multiple columns individually. In the example below, we have selected the column for DATE, along with the column representing TOTAL NUMBER OF FILES SENT.

If you would like a graphical breakdown of the total bytes sent for a particular timeframe, you would select the DATE column and the SUM(file_size) column. The SUM(file_size) field represents the total data transferred in bytes.

	A	B	C	D
1	DATE(date_transmitted)	SUM(file_size)	count(DATE(date_transmitted))	
2	5/1/2007	80067026461	37741	
3	5/2/2007	73774069708	35160	
4	5/3/2007	73580223764	35321	
5	5/4/2007	74369839358	34875	
6	5/5/2007	74185093584	35192	
7	5/6/2007	69197866109	34350	
8	5/7/2007	67959925986	33513	
9	5/8/2007	33791322075	16719	
10				
11				
12				

Microsoft Excel Chart Wizard – Selecting Chart Data Areas (multiple, separate columns)

22. The Source Data window will appear, showing a mini-representation of the data you have selected. **Be sure to include the column headers, as these will represent the chart titles and data labels.** Click the "Next" button to proceed.
23. Click the "Finish" button to paste your chart into your worksheet.



Example of a Completed "Total Files Transferred By Date" report within an Excel Spreadsheet