

1 WHAT IS FILE QUERY?

File Query is a program that allows you to quickly and easily view, query, analyze and modify flat files (delimited or fixed-width) directly, as though they were in a database.

2 THE INTERFACE

The screenshot displays the File Query application window titled "Customer_Sample.txt - File Query (2.10)". The interface includes a menu bar (File, Run, Help, Debug), a toolbar with icons for layout, find, options, and set fields, and a status bar showing "00:00:00 Current Line", "LPH:", and "Time Remaining:". The main workspace is divided into several sections:

- 1**: Layout options (D, F, T, Options, Set Fields).
- 2**: Run button and status indicators.
- 3**: Input and Output file paths: "C:\Users\Jeff\OneDrive\Documents\AgileData Software\Documentation\File Query\Customer_Sample.txt" and "C:\Users\Jeff\OneDrive\Documents\AgileData Software\Documentation\File Query\Customer_Sample_Out.txt".
- 4**: Field Names list:


```
0: CUSTOMER_ID
1: CUSTOMER_TYPE
2: CUSTOMER_STATUS
3: RESPONSE_CODE
4: FIRST_NAME
5: LAST_NAME
6: GENDER
7: ADDRESS
8: CITY
9: STATE
10: ZIP_CODE
11: LAST_TRANSACTION_DATE
```
- 5**: Query editor with a sample query:


```
var query = {
  from line in inputFile
  select line
};
```
- 6**: File Preview window showing a table of customer data with columns: CUSTOMER_ID, CUSTOMER_TYPE, CUSTOMER_STATUS, RESPONSE_CODE, FIRST_NAME, LAST_NAME, GENDER, ADDRESS, CITY, STATE, ZIP_CODE, LAST_TRANSACTION_DATE. The data is displayed in a grid format.
- 7**: Status bar showing "Line Count: ~1,008", "Records Returned:", "Cur Line: 1", "Cur Char: 1", "ASCII Value: 67/43H", "Selection Length: 0", and "EOL: CRLF".

- 1. Input File Layout** – Displays, and lets you edit how the input file is interpreted. It can be either a fixed-width or delimited file, have a header, and be quoted. The two options allow you to trim fields on input or upper-case all text, and the 'Set Fields' button allows you to update field layout information.
- 2. Query Control** – Allows you to start and stop queries, set if the results should be written to a file, and if you should append to any existing file. It also displays the current status of a query while it is running, such as how long it's been running for, what line of the input file it is on, and how

many lines per hour it is currently processing. For running queries, you can also use F5 to start them, and Ctrl+Q to stop them.

3. **File Paths** – Shows the current input file path, and the location that any output file would be written to.
4. **Field List** – Displays a list of fields in the file. The buttons below let you change the sort order, or copy the list to the clipboard.
5. **Query Editor** – Where you enter a query to run against the file. Most of the time, you would use the built in functions (top-right of the query window) instead of writing them yourself. Below the query are built in templates, separated into various categories. The 'Templates' category contains full queries, and the other categories contain specific functionality that can be used to build or alter new queries.
6. **File Preview/Results** – Displays a preview of the file, as well as the results of a query. The two buttons in the top right allow you to refresh the preview and spread the fields out into even, easy to read, columns. The spreading of fields works on both the preview and results view.
7. **Status Bar** – Displays a line count (estimate for delimited files, and exact for fixed-width files), the number of records returned from a query, information about the current character of the preview or results window, and the type of end-of-line character used in the file.

3 USING FILE QUERY

Everything in File Query is based off of a query (go figure), which uses a custom syntax. It is similar to SQL in what it does, but works a little differently. Luckily, File Query has built in functions that write and run them for you, so most people will never have to worry about them. At the top-right of the query box, you will see menu options for the different function types (Find, Modify, Analyze/Validate), and within those contain the actual functions.

Below, are examples of how to use File Query for its four main functions (View, Query, Analyze, Modify), but is just enough to get you started. After going through the examples, and trying them out yourself, you should go to www.agiledatasoftware.com/support.html to find more information about what you can do with File Query, and also a form that you can use to ask us any questions you still have. We would love to hear from you.

3.1 VIEWING A FILE

The first thing you will often want to do with a file is just get a feel for what it looks like, as far as the format, and the type of data that it contains. When File Query opens a file, it only reads in the first 100

lines at first, which allows you to open huge files that normal text editors, such as Notepad, can't even open. For delimited files, it gets even better, as it will automatically determine the delimiter type, if it has a header and if it is quoted. It will then display a list of all fields, and allow you to sort them, or copy them to the clipboard.

One of the most popular features is the ability to hit the Normalize button at the top-right of the preview (or hit Ctrl+N), and spread out delimited fields into even and readable columns. Go ahead and try it.

3.2 QUERYING A FILE

When you are looking for something specific within a file, there are multiple queries to help you find it. We will just look at one of them for now, but be sure to check out the others yourself. Each function has an Info section at the bottom that explains what it does, and how it works, so you shouldn't have any trouble.

3.2.1 Find Lines Where...

Under the Find menu, you will see the 'Find Lines Where...' function, which is basically a Swiss-Army knife for finding certain lines within a file. It allows you to pick up to three criteria that a line must meet to be selected, and you can set it so that the line must meet all of them, or just one of them.

Find Lines Where...

Field Names

- 0: CUSTOMER_ID
- 1: CUSTOMER_TYPE
- 2: CUSTOMER_STATUS
- 3: RESPONSE_CODE
- 4: FIRST_NAME
- 5: LAST_NAME
- 6: GENDER
- 7: ADDRESS
- 8: CITY
- 9: STATE
- 10: ZIP_CODE
- 11: LAST_TRANSACTION_DATE

Settings

Select all lines where:

field	Field Modifier (optional)	Comparison	Comparison Value	Reverse
9: STATE		In	MA,NH,CT,RI,VT,ME	<input type="checkbox"/>
and 8: CITY	.Length	>	50	<input type="checkbox"/>
and 2: CUSTOMER_STATUS		==	ACTIVE	<input type="checkbox"/>

From each record found, select entire line fields 0, 1, 2 Include line number

Info

This function allows you to select only records that meet your criteria, and from those records, you can select either the entire line or just certain fields. The 'Reverse' checkbox allows you to switch the meaning of each comparison, so for example, checking that field 0 was empty would actually check that it was not empty.

Sort by Index Name

Only read in 1000 Run Query Done Cancel

Here are three examples of things that you can search for, although you can combine things together to be able to search for anything.

- **Find a Specific Value in a Field:** Select the field you want to search in, leave the Comparison option at the default equal sign, and then enter is the value you want to search for in the Comparison Value box.

- **Find Values Greater Than a Certain Length:** Select the field to check, select '.Length' as the Field Modifier, select '>' as the Comparison, and then enter the length you want to compare to.
- **Find Empty Values:** Select the field to check, and select 'IsEmpty' as the Comparison. You can also check the Reverse option to find values that are **NOT** empty.

3.3 ANALYZING A FILE

Instead of looking for specific records in a file, you can do some analysis to find out the type of data you have, as well as verify certain criteria. Here are some common examples.

3.3.1 Find Duplicate Values

Under the Analyze/Validate menu is the Duplicate Values function, where you can select up to three fields that will be combined together, and look for any duplicates across them. You can even specify that you only want to use the first few characters of any of the fields, such as the first 5 digits of a zip code that may or may not have the zip+4 after it.

3.3.2 Frequency Report

Underneath the Find Duplicates function is the Frequency Report. This report allows you to analyze all fields at once (or only some of them), and get information about the values contained in each of them. This data includes the following information:

- # of distinct values
- # of empty values
- Min and max length of the values
- Character types used (L=Letters, D=Digits, W=Whitespace, and special characters represent themselves)
- Distinct values, and how many times each occurs

Some things to note about this report: The option about how many unique values to track stops the report from keeping a list of the individual unique values, and how many times they occur, for fields that have too many unique values (such as an ID field), but all other aggregate information is still reported, such as min/max length and character types. Also, if you select the Delimited Format option, it will leave out the list of distinct values, and will show the aggregate info for each field on a single line. This is great for getting a high-level overview of what the data in the file looks like, but be sure to use the normalize

button to spread the columns out, so that it is readable, or just copy and paste it into Excel if you want to put it in a separate report.

3.4 MODIFYING A FILE

It is important to understand how File Query works when it comes to modifying files. First, is that File Query will never actually modify a file directly, but will instead write out a modified version to a new file. This ensures that you never ruin the original data, and can always go back if you change your mind or make a mistake. Also, any query can be written out to a file by checking the 'Write to File' option below the query entry box, so you aren't limited to just what is under the Modify menu. If you want a version of a file with only records from MA, just select only those records with the 'Find Lines Where...' function, and write it out.

For simple tweaks though, you will most likely find what you need under the Modify menu.

3.4.1 Updating a Field

The Update Fields function allows you to modify up to three fields by either setting them to a specific value, setting it to a substring of itself, uppercasing it, replacing text within it, or setting it to one of two values, based on its current value. You can also update the fields for all lines, or only those that match certain criteria.

A simple example could be setting the SHIPPED field to 'Y', but only if the state is 'NH'. To do this, you would select the field you want to update (SHIPPED) from the field list, leave the Update Type option to 'Set Value', and then enter 'Y' (without the quotes) as the Update Value. This will set it to Y for all lines, so to do it only for NH records, you would then select the 'For all lines where...' option, and use it just like the 'Find Lines Where...' function, by selecting STATE as the field, and 'NH' (without the quotes) as the comparison value.

Be sure to check the 'Write to File' option as well if you want to create a new updated file.

3.4.2 Modifying the Layout

The Modify Layout function is a simple way to adjust the layout of your file by adding, removing or re-arranging fields, as well as changing the delimiter type, and adjusting the header. It is very straightforward, but one helpful tip is that if you are adding a new field, and want it to have a default value, instead of being empty, just enter the field name as 'FIELD_NAME=DEFAULT_VALUE', such as 'ENROLLED=N'. Also, re-arranging fields is done by dragging the fields around in the fields list, and if you want to select multiple fields at once, just hold the Shift or Ctrl key while click on the fields.

4 LEARNING MORE

This guide just glossed over a few of the more common functions found in File Query. There are many more for you to use, but they should be very straightforward to figure out on your own. If you want a more detailed introduction to using File Query, just go to the support page of www.agiledatasoftware.com, and look at the Getting Started video series there, or send us a message if you still have any questions.