Data Mining

Help Documentation

Data Mining

Data Mining Overview

While the default reports available in SmarterStats provide a general overview of the various statistics extracted from your log files, data mining allows site administrators and users to drill-down and analyze statistical data on a deeper level. These reports are an invaluable resource for tracking customers through your ordering process, for seeing which marketing campaigns are providing the highest return on your investment(s) or for a variety of other tasks requiring a large-scale, detailed analysis of website traffic.

To reduce the complexity often associated with data mining, you can associate a Query Item, and the data returned, to an intuitive question and answer format. For example, by asking the questions, "What are the Referrers of this file?" or "What Paths were taken through a particular Page?" Using this simple method, businesses can get the insight they need to evaluate things like the effectiveness of an online marketing or advertising campaign.

To access data mining reports, click the Data Mining icon. You'll be presented with the following:

Query Items

The type of data returned in Data Mining is dependent on what it is you want to dig down into. By default, SmarterStats offers the following Query Types in the Navigation Pane:

Daily Activity

This query allows you to select a single page or resource contained within your website and see the daily activity over time. This type of mining is good to identify trends in the activity of that resource and to see if people are accessing the file more or less than they used to.

Entry Pages

This query allows you to select a single page contained within your website and see the top pages used to access the page. This is a particularly useful form of data mining that allows you to track where people are coming from when they access a specific page. For example, if a page is specified in an advertising campaign or marketing campaign, users can see exactly how people are accessing that page and verify any information that is given by the particular site that contains the advertising.

Exit Pages

This query allows you to select a single page contained within your website and see the last pages

people viewed before exiting the website. This is useful in order to track how far users get in a process (such as placing an order) and where they are exiting your site.

IP Addresses

This query allows you to track which visitors (IP addresses) most often requested a file on your website. When used in conjunction with other data mining queries, this report is helpful in understanding the behavior of visitors that request certain files.

Paths through Page

This query allow you to track the most common paths taken for visitors that request a specific file. This report is useful in understanding how visitors are accessing various files on your website.

Query Items

This query allows you to view the various query string items that are passed to a particular page of your website. A typical example is the SmarterStats help system. Whenever anyone clicks on the "Help" graphic or a "What is this?" link within a specific report, a query string is passed from the SmarterStats application to the SmarterTools Online Help. This query string relates to the particular page or report the user is seeking help with. This report shows which individual items were passed to generate the page. Note: This report differs from the Query Strings report in that it separates each query string item and reports them individually. For example, assuming you have a hit that looks like /help/default.aspx?section=Administrator&page=Getting+Started. The Query Items report will show two lines: One for section=Administrator and one for page=Getting+Started. Conversely, the Query Strings report will show one line that looks like section=Administrator&page=Getting+Started.

Query Strings

This query allows you to view the various query strings that are passed to a particular page of your website. A typical example is the SmarterStats help system. Whenever anyone clicks on the "Help" graphic or a "What is this?" link within a specific report, a query string is passed from the SmarterStats application to the SmarterTools Online Help. This query string relates to the particular page or report the user is seeking help with. This report shows which query string was passed to generate the page. Note: This report differs from the Query Items report in that it does not separate query string items. For example, assuming you have a hit that looks like /help/default.aspx?section=Administrator&page=Getting+Started. The Query Items report will show two lines: One for section=Administrator and one for page=Getting+Started. Conversely, the Query Strings report will show one line that looks like section=Administrator&page=Getting+Started .

Referrers

This query is useful for marketing or to detect image leeching. For example, if you have an image that you have exposed on other websites that resides on your Web server, the URL it was viewed on is recorded every time a browser views your image. This report will list the top URLs or "referrals" in which your image was viewed. If you have a nice graphic, chances are that someone else will want to use it on their website. If they link your image path to their website, this report will list the top URLs that the image or graphic is viewed from.

Referring Sites

This query reports the original source of visitors that came into your site that eventually hit the page or file you choose. This can be most useful for marketing purposes. For example, if you have a "Thank You" page after a person orders your product, you can identify the original sources of visitors that reached that page, effectively discovering where your orders came from.

Usernames

This query allows you to track which authenticated users hit a specific page of your site. This report is useful in understanding the behavior of specific visitors on your website.

Visitor Bandwidth

This query allows users to track the bandwidth usage, total bandwidth usage, and the IP address of requests for specific files you have placed on your website. This report is most useful in detecting web site abuse. Typically, you will identify a high-bandwidth file from Top Files or a related report item, then start data mining on it to discover if there are a few IP addresses that are attempting to use your bandwidth by repeatedly downloading files. Those IP addresses can then be metered or denied access on your Web server.

Visits Before Hit

This query allows you to select a single page contained within your website and see the number of times a person visited the site prior to reaching and viewing the page selected. The report itself shows the IP address of the visitor, the initial entry page for that visitor before they viewed the page selected, and the number of visits before the page selected was viewed.

World Cities

This query allows you to identify the geographic regions that accessed a specific file on your website.

Top Files from IP Address

This query allows you to track the files most commonly requested by a particular IP address. This report is useful in understanding the behavior of specific visitors on your website.

Top Paths from IP Address

This query allows you to track the most common paths taken by an IP address during its visits to your website. This report is useful in understanding the behavior of specific visitors on your website.

Report Options

- Filename - The file to be data mined. To browse for a specific file, click the folder icon to the right.
- Start and End Date - The time frame on which the data mining report statistics are generated.
- Chart - The chart type displayed in the data mining report.
- Group By - When available, the timeframe you want the data returned to be grouped by. (E.g., Weekly)
- Sort - The column that will be used to sort the data in the report. (E.g., Views)
- Rows - The number of rows in the data mining report.
- Filter Set - If a filter set exists, this allows you to seledt the filter set that should be applied to the data mining report.

Where applicable, you can also perform data mining by selecting it from the dropdown arrow next to items returned in a report.