Google® Search Tips and Techniques for SAS® Users

Kirk Paul Lafler, Software Intelligence Corporation, Spring Valley, California
Charles Edwin Shipp, Consider Consulting, San Pedro, California

Abstract
Google® (www.google.com) is the world’s most popular and widely-used search engine. As the premier search tool on the Internet today, SAS® users frequently need to identify and locate SAS-related content wherever and in whatever form it resides. This paper provides insights into how Google works and illustrates numerous search tips and techniques for finding articles of interest, reference works, information tools, directories, PDFs, images, current news stories, user groups, and more to get search results quickly and easily.

Introduction
As the world’s information grows to astronomical levels the world’s largest search engine, Google, and its proprietary software, organizes this information and makes it useful and accessible to everyone. Google users are well aware of the speed, accuracy, and reliability that a Google search provides. Because of this, SAS users frequently turn to Google for their search needs because of its ability to find the information they want, when they want it, providing them with the speed, accuracy, and organization of the searched results. In this paper, the authors take you on a journey into the world of Google by starting with the Google user interface, showing you how Google works, exploring various search techniques, all the while sharing an assortment of tips and techniques that SAS users can use to achieve better searches and better results.

SAS Specific Websites
Users have several website choices to use when searching for SAS specific and SAS-related content. For starters, SAS Institute’s technical support website, support.sas.com, provides users with access to the world’s best technical support content, 24-7. Users can access a comprehensive Knowledge Base containing information related to products & solutions, system requirements, install center, third-party software reference, documentation, papers, samples & SAS notes, and focus areas; Support with technical issues such as submit a problem, update a problem, check problem status, license assistance, manage my software account, and downloads & hot fixes; Training & Bookstore with access to the bookstore, training, certification, SAS global academic program, and SAS on-demand; and Community with access to user groups, SAS talks, events, SASware ballot, e-newsletters, author with SAS, RSS & blogs, and support communities, see Figure 1.

Figure 1. support.sas.com – SAS Institute’s Technical Support website
Another website available to users for accessing a repository of more than 27 thousand published “white” papers consisting in PDF format is found at www.lexjansen.com. SAS specific and SAS-related content from SAS Global Forum (SGF) and SUGI international conferences; MWSUG, NESUG, PNWSUG, SCSUG, SESUG, and WUSS regional conferences; and PharmaSUG, PhUSE, and CDISC special-interest conferences, are available to users to search anytime and anywhere, see Figure 2.

And finally, a “virtual” website that SAS users can use to search for SAS specific and SAS-related content is available at www.sascommunity.org. sasCommunity.org is a collaborative online SAS Community for SAS Users worldwide, see Figure 3.
The Google User Interface

When the SAS specific websites are not sufficient for all your research needs, Google’s “free” and easy-to-use Basic Internet search service begins with a very familiar user interface (or home page). Using a web browser such as Google Chrome®, Mozilla Firefox®, Internet Explorer®, or Safari®, the web address, www.Google.com, is entered as shown in Figure 4. By entering a keyword (or phrase) in the search box (section 1) and clicking the “Google Search” button (section 2), a basic user-initiated search can be requested. In addition to using the Google home page to search relevant results on the World Wide Web, users are also able to perform specific searches (i.e., You, Search, Images, Maps, Play, YouTube, News, Gmail, Documents, Calendar, and More) by clicking the links located at the top of the Google page (section 3).

![Figure 4. The Google User Interface](image)

How Google Works

So, how does Google work? You’re definitely not the first to ask this question, and you won’t be the last. Our explanation of how Google works will be separated into two distinct phases: 1) Google’s web crawling, extraction and indexing process, and 2) Google’s query processor process. The web crawling, extraction and indexing process, as illustrated in Figure 5, shows Google’s automated web crawler, or a computer program (aka spider or Googlebot), browsing, extracting, and indexing (organizing) content from the World Wide Web by thousands of Google computers. Essentially, tens of thousands of Googlebots crawl (search) broad expanses of the Internet harvesting web page links creating a list of links. The list of links are organized, indexed and then stored in databases.

![Figure 5. Web Crawling, Extraction, Organizing, and Indexing Process](image)
The query processor process, as depicted in Figure 6, shows that as a query is submitted, the Google web servers immediately send the query to the index servers to determine the pages that contain the word(s) that match one or more query terms. The query is then sent to the doc servers where the stored content is retrieved, and the search results immediately returned to the Google user.

Measuring the Importance of Web Content with Google’s PageRank®

What makes Google search the most popular and widely used search application in the world? There are many factors, but one essential component is the creation of the ingenious ranking of web pages, links and content, known as PageRank®, developed by founders Lawrence Page and Sergey Brin. The PageRank algorithm ranks (or scores) web content with the greatest importance so that content is moved to the top of the user’s search results allowing the Google web, index, and doc servers to return relevant and validated search results quickly. For example, a Google-search on [pagerank algorithm code] returns 32 million results, ranked with the most popular and important results appearing first in less than a second.

Currently, Stanford University holds the patent and Google has exclusive license rights to the proprietary PageRank algorithm. Although the exact number and type of parameters used in the ranking algorithm (originally dubbed “Backrub”) is probably only known by a select few, we are able to peek under the hood, so to speak, by reading an important published paper (1998) detailing the page ranking formula, by Sergey Brin and Lawrence Page, when Google was just a university project. The formula for calculating PageRank and the explanation of its parameters is shown in Table 1, below.

\[
PR(A) = (1-d) + d \left( \frac{PR(T1)}{C(T1)} + \ldots + \frac{PR(Tn)}{C(Tn)} \right)
\]

Where \(PR(A)\) is the PageRank of Page A.

\(D\) is a dampening factor. Nominally this is set to 0.85.

\(PR(T1)\) is the PageRank of a site pointing to Page A.

\(C(T1)\) is the number of links off that page.

\(PR(Tn) / C(Tn)\) means we do that for each page pointing to Page A.

Source:

Table 1. Formula for calculating PageRank, by Sergey Brin and Lawrence Page (1998)
An essential feature of Google’s ranking algorithm applies greater importance to web content that contains page links from other pages. The algorithm also places greater importance to web content when it contains important links so that any links to other web pages also, by inheritance, become important.

In their in-depth and landmark papers, *PageRank Uncovered*, written by Chris Ridings and Mike Shishigin, and *How Google Ranks Web Pages* by Brian White, the basic features of Google’s proprietary PageRank algorithm and how it ranks web content follows.

1. Find web content matching the keywords of the search.
2. Using the PageRank algorithm and its more than one hundred parameters, web content is ranked.
3. Compute the ranks once each month.
4. Return a list of the relevant pages using the current month’s rankings.

**Better Searches = Better Results**

Are you a savvy Google search user? Do you know search tips and techniques to effectively find the content you’re looking for on the Internet? For many, the importance of knowing how to conduct successful searches is not only important, it may be an essential activity required by their job. Because the Google search engine adheres to rules and processes the built-in algorithms make every attempt to interpret search requests while delivering the “best” results possible; in the end, the derived results are only as good as the search terms that are provided. To alleviate the many challenges of finding the right combination of keywords or phrases along with the frustration associated with unsuccessful searches and massive listings, the following tips and techniques are designed to achieve better searches and better results.

**Basic Search Tips and Techniques**

Understanding basic search techniques gives you incredible power to find what you are looking for quickly and easily. In addition to this, you may actually find content you didn’t plan for, but adds value to your query results. Table 2 below, presents a number of basic search tips and techniques you may find useful the next time you prepare a search query.

<table>
<thead>
<tr>
<th>Tip/Technique</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Keep it Short</strong></td>
<td>Limit the number of search words specified for greater accuracy in the generation of search results.</td>
<td>sascommunity</td>
</tr>
<tr>
<td><strong>Be Specific</strong></td>
<td>When entering your search, be as specific as possible. Avoid generic (or vague) references when entering search words.</td>
<td>BOOK proc sql</td>
</tr>
<tr>
<td><strong>Use Quotes</strong></td>
<td>Search the words in the exact order specified.</td>
<td>“proc sql” &lt;or&gt; “sas sql”</td>
</tr>
<tr>
<td><strong>Search a Website</strong></td>
<td>Have your search results come from a specific website.</td>
<td>hash SITE:www.sas.com</td>
</tr>
<tr>
<td><strong>Wildcard (*)</strong></td>
<td>The asterisk (*) is used as a placeholder (or wildcard) for unknown words.</td>
<td>sas blogs *</td>
</tr>
<tr>
<td>**OR &lt;or&gt;</td>
<td>**</td>
<td>The OR (in all CAPS) tells Google to search either one of several words. The pipe symbol (</td>
</tr>
<tr>
<td><strong>AND</strong></td>
<td>The AND is the default logical operator that Google uses between two or more words to search web content, and is therefore not required. Specifying ‘AND’ in UPPERCASE produces the same results.</td>
<td>sas AND “sas sql”</td>
</tr>
<tr>
<td><strong>– (Minus sign)</strong></td>
<td>A word not wanted in a search query can be excluded by specifying a – (minus sign) before the word. Note: A blank space should precede a – (minus sign) to avoid confusing it with a hyphenated word.</td>
<td>sas -airlines</td>
</tr>
<tr>
<td><strong>Excluding Words</strong></td>
<td>More than one word can be excluded from your search query by specifying a – (minus sign) before each word.</td>
<td>sas –airlines –shoes –military</td>
</tr>
<tr>
<td><strong>Computations</strong></td>
<td>Google can perform basic arithmetic computations.</td>
<td>77 x 119 &lt;or&gt; pi x 7</td>
</tr>
</tbody>
</table>

*Table 2. Basic Search Tips and Techniques*
“Powerful” Specialized Google Operators

Google provides a number of specialized operators to help you with your search queries by finding information about a specific book, population number, investment fund, movies, public stock, unemployment rate, weather, or website; identifying and displaying information that Google has collected about backlinks (or incoming links) for a specific website; and display maps about a specific country, state, city, or location, as shown in Table 3.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>AREA CODE</td>
<td>Display the geographical location for any three-digit area code.</td>
<td>AREA CODE 619 &lt;or&gt; AREA CODE 310</td>
</tr>
<tr>
<td>BOOK</td>
<td>Search and display book-related information.</td>
<td>BOOK proc sql &lt;or&gt; BOOK sas sql</td>
</tr>
<tr>
<td>DEFINE</td>
<td>Display a definition for a specific word or phrase.</td>
<td>DEFINE quasar &lt;or&gt; DEFINE Miami</td>
</tr>
<tr>
<td>INFO</td>
<td>Display information that Google has collected for a website.</td>
<td>INFO <a href="http://www.sas.com">www.sas.com</a></td>
</tr>
<tr>
<td>LINK</td>
<td>Display backlinks (or incoming links) for a specific website or web page that is received from another website.</td>
<td>LINK <a href="http://www.sas.com">www.sas.com</a></td>
</tr>
<tr>
<td>MAP</td>
<td>Display a map of a specific country, state, city or location.</td>
<td>San Diego map</td>
</tr>
<tr>
<td>MOVIE</td>
<td>Search and display where a movie is currently playing and a description of all movies currently playing in a specific location.</td>
<td>MOVIE bourne &lt;or&gt; MOVIE 91978 &lt;or&gt; MOVIE spring valley</td>
</tr>
<tr>
<td>POPULATION</td>
<td>Display the population of a U.S. state or county.</td>
<td>POPULATION bourne San Diego</td>
</tr>
<tr>
<td>SITE</td>
<td>Display the number of indexed pages for a specific website.</td>
<td>SITE <a href="http://www.sas.com">www.sas.com</a></td>
</tr>
<tr>
<td>STOCK</td>
<td>Display the market data for a specific company’s stock or fund.</td>
<td>STOCK goog &lt;or&gt; STOCK siri</td>
</tr>
<tr>
<td>TIME</td>
<td>Display the current time in a city.</td>
<td>TIME Honolulu</td>
</tr>
<tr>
<td>UNEMPLOYMENT RATE</td>
<td>Display unemployment rate trends of a U.S. state, county or zipcode.</td>
<td>UNEMPLOYMENT RATE San Diego &lt;or&gt; UNEMPLOYMENT RATE 91978</td>
</tr>
<tr>
<td>WEATHER</td>
<td>Display the weather conditions, temperature, humidity, wind, and forecast for many cities or zipcodes.</td>
<td>WEATHER San Diego</td>
</tr>
<tr>
<td>~ word</td>
<td>Search and display a synonym or similar word.</td>
<td>~ statistician</td>
</tr>
</tbody>
</table>

Table 3. Specialized Google Operators

Advanced Search Tips and Techniques

Google offers advanced search tips and techniques which are worth knowing for searching specific file types, results between two values, within a specific date range, and from related websites, as shown in Table 4.

<table>
<thead>
<tr>
<th>Tip/Technique</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>FILETYPE: (File Type)</td>
<td>Google automatically displays all available file types by default but can be told to display specific file types using the FILETYPE: operator. The values for FILETYPE: include: DOC, DWF, KML, KMZ, PDF, PPT, PS, RTF, SWF, and XLS.</td>
<td>sas FILETYPE:PDF</td>
</tr>
<tr>
<td>$nnn…$nn (Price Range)</td>
<td>Google can show search results for PROC SQL books priced $19 to $50 by using three dots between the numeric values.</td>
<td>proc sql book $19 . . . $50</td>
</tr>
<tr>
<td>DATERANGE:yyyymm (Date Search)</td>
<td>Google can show search results containing information and news from a specified date using the DATERANGE: operator. Note: The date range value is expressed as DATERANGE:YYYYMM where YYYY=year and MM=month.</td>
<td>‘sas hash’ DATERANGE:201206</td>
</tr>
<tr>
<td>RELATED: (Sites that are Similar)</td>
<td>Google can show search results from websites that are similar (in the opinion of Google) by specifying the RELATED: operator.</td>
<td>sas RELATED:www.sas.com</td>
</tr>
</tbody>
</table>

Table 4. Advanced Search Tips and Techniques

Conclusion

As the world’s information continues to grow to astronomical levels, the world’s largest search engine, Google, and its proprietary software, organizes this information and makes it useful and accessible to everyone. This paper offers numerous tips and techniques which, if used, should help you take advantage of the speed, accuracy, and reliability that a Google search provides. As a result, SAS users frequently turn to Google because of its ability to find the information they want, when they want it, while providing them with the speed, accuracy, and organization of the searched results.
References

Acknowledgments
The authors would like to thank Jason Brinkley and Joe Kelley, Planning, Support, and Administration Section Chairs, for accepting our abstract and paper; William E. Benjamin Jr., SESUG 2015 Academic Chair, Venita DePuy, SESUG 2015 Operations Chair, the SouthEast SAS Users Group (SESUG) Executive Board, and SAS Institute for organizing and supporting a great conference!

Trademark Citations
SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Google is the registered trademark of Google Inc., Mountain View, CA, USA. Other brand and product names are trademarks of their respective companies.
About the Authors

Kirk Paul Lafler has been programming in SAS since 1979, and is consultant and founder of Software Intelligence Corporation. He is a SAS Certified Professional, provider of IT consulting services, trainer to SAS users around the world, mentor, and sasCommunity.org emeritus Advisory Board member. As the author of six books including Google® Search Complete! (Odyssey Press. 2014) and PROC SQL: Beyond the Basics Using SAS, Second Edition (SAS Press. 2013); Kirk has written more than five hundred papers and articles, been an Invited speaker and trainer at five hundred-plus SAS international, regional, special-interest, local, and in-house user group conferences and meetings, and is the recipient of 23 “Best” contributed paper, hands-on workshop (HOW), and poster awards.

Charles Edwin Shipp is a programmer, consultant and author, and has been using the SAS and JMP software since 1980. He is credited in the original JMP manual for his roles in the early days. He has written more than one hundred papers and has been an invited speaker at more than one hundred International, regional, local, and special-interest groups. He is the recipient of 13 “Best” contributed paper and poster awards. Charlie is the co-author of three books including Google® Search Complete! (Odyssey Press. 2014) and Quick Results with SAS/GRAPH Software. Currently, Charlie is involved as an eBook author, App developer for Apple iPad, sasCommunity.org Advisory Board member, consultant for 4Life, AdvoCare, Genesis Pure, Melaleuca, Trivani Foundation International, and JMP consulting.

Comments and suggestions can be sent to:

Kirk Paul Lafler  
Senior SAS® Consultant, Application Developer, Data Scientist, Trainer and Author  
Software Intelligence Corporation  
E-mail: KirkLafler@cs.com  
LinkedIn: http://www.linkedin.com/in/KirkPaulLafler  
Twitter: @sasNerd  
~ ~ ~ ~ ~

Charles Edwin Shipp  
Senior SAS® and JMP® Consultant, Programmer, eBook Author, Trainer and Author  
Consider Consulting Corporation  
E-mail: CharlieShipp@aol.com  
LinkedIn: http://www.linkedin.com/in/CharlieShipp  
Twitter: @ShippAhoy