

Let me 'Google' that for you - Google Labs Code Search

James Bidwell, Amgen Ltd, Uxbridge, UK

ABSTRACT

Google Labs Code Search is a free beta product from Google which debuted in Google Labs in 2006, allowing web users to search for open-source code on the Internet. Despite opinions to the contrary, this product can be utilised very effectively for finding SAS code snippets and syntax examples.

Depending on how you personally learn, remember or improve your SAS knowledge, Google Labs Code Search should at least be tried once, and may turn out to be your number one SAS code resource.

This paper will present the advantages, tips, tricks and annoyances of using the search engine on a daily basis for SAS code development.

INTRODUCTION

Google Labs is a website demonstrating and testing new Google projects. Widely used products which have graduated from the labs in the past include Google Maps, Google Suggest and Google Desktop.

Google Labs Code Search is a free beta product from Google, allowing web users to search for open-source code on the Internet. This product can be used very effectively for finding SAS code snippets and syntax examples.

Depending on how you personally learn, remember or improve your SAS knowledge, Google Labs Code Search should at least be tried once, and may turn out to be your number one SAS resource. The power of the search technology can identify certain syntax (e.g. % & =) that would otherwise be ignored when using other online resources such as the SAS OnlineDoc Manual and the standard Google Search. Subtle differences in the uses of syntax can be explored when the results from your search show many SAS examples on one easy-to-read page.

Do you always forget the FORMAT syntax for the SELECT statement under PROC SQL? Or the order of the variables in a CATX or CATS function? Simple...use the Google Labs Code Search and the results will quickly show you the solution.

This paper will discuss the advantages, highlights, tips, tricks and annoyances of using the search engine on a daily basis for SAS code development.

Google Labs Code Search, when fully developed and tested, will be one Google product to keep an eye on.

HOW DO YOU LEARN, REMEMBER OR IMPROVE SAS?

SAS programmers learn their trade using a variety of methods, but we should also strive to improve our SAS skills. A few of the resources some might use are listed below.

- SAS Course Notes
- SAS Online Manual
- Personal Notes
- Books
- Stored Code Library on personal computer
- Internet

A combination of resources is often the most effective way, and after reading this paper, giving the new methods a trial, I would hope that Google Labs Code Search will become a valuable resource too.

GOOGLE LABS CODE SEARCH IN USE...

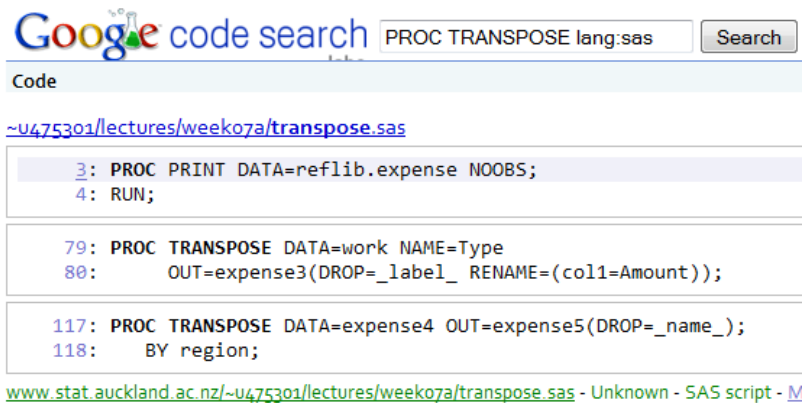
The interface will be familiar with anybody who has searched for items on the internet before. In a similar manner to the standard search engine, there is a useful 'Advanced Search' option. You will have to select SAS as the default language either in the 'Advanced Search' option or by typing "lang:SAS" as an additional search parameter.

PhUSE 2010



The image shows the Google code search labs interface. At the top is the Google logo with 'code search labs' underneath. Below the logo is a search bar containing the text 'PROCTRANSPOSE lang:SAS'. To the right of the search bar is a 'Search Code' button and a link to 'Advanced Code Search'. Below the search bar is a green link that says 'Search public source code.' Underneath this is a section titled 'Syntax and Examples (more about regexp syntax)'. This section contains two rows: the first row is for 'regexp' with the example 'go(2)gle hello.\world ^int printk'; the second row is for 'exact string' with the example '"compiler happy"'.

After submitting your search, the results window appears with many snippets of code from SAS programs across the internet, mainly focusing on the programs named .SAS. At present the Search Engine does not search inside PDF files, for this, the main Google Search Engine should be used. The results are mainly from Educational resources, as industry code is not usually published.



The image shows the search results for 'PROC TRANSPOSE lang:sas'. The results are displayed in a list of code snippets. The first snippet is from a file named 'transpose.sas' and contains the following code:

```
3: PROC PRINT DATA=reflib.expense NOOBS;
4: RUN;
```

 The second snippet is from a file named 'transpose.sas' and contains the following code:

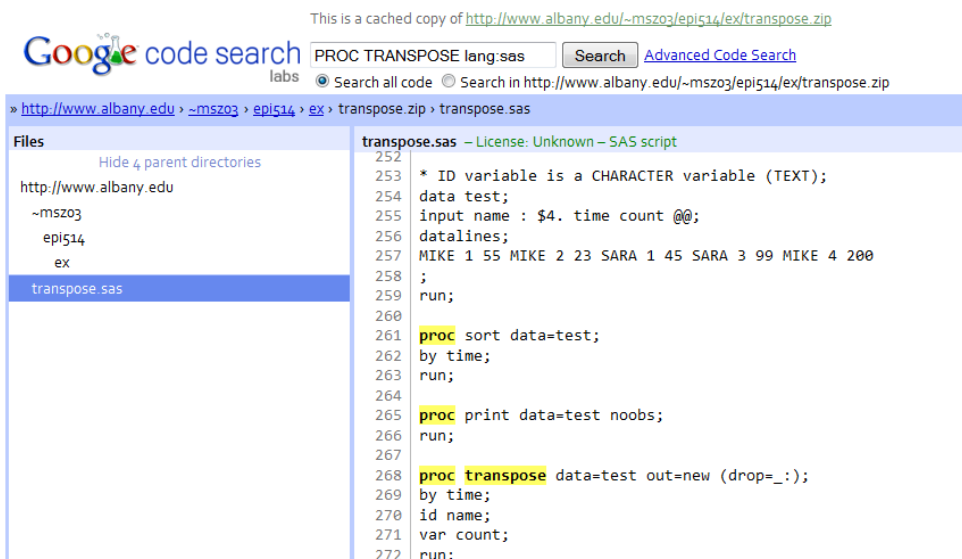
```
79: PROC TRANSPOSE DATA=work NAME=Type
80:   OUT=expense3(DROP=_label_ RENAME=(col1=Amount));
```

 The third snippet is from a file named 'transpose.sas' and contains the following code:

```
117: PROC TRANSPOSE DATA=expense4 OUT=expense5(DROP=_name_);
118:   BY region;
```

 Below the snippets is a link to the source file: www.stat.auckland.ac.nz/~u475301/lectures/week07a/transpose.sas - Unknown - SAS script - M

The search engine gives the capability to “click into” any of the code on display. This will open up the code, highlighting the search parameters, and will also give a structure diagram of the directory if you wanted to check the .LOG and .LST file also.



The image shows the detailed search results for 'PROC TRANSPOSE lang:sas'. At the top, it says 'This is a cached copy of http://www.albany.edu/~msz03/epi514/ex/transpose.zip'. Below this is the Google code search labs interface. The search bar contains 'PROC TRANSPOSE lang:sas'. To the right of the search bar is a 'Search' button and a link to 'Advanced Code Search'. Below the search bar is a section titled 'Files'. This section contains a list of files: 'http://www.albany.edu', '~msz03', 'epi514', 'ex', and 'transpose.sas'. The file 'transpose.sas' is selected. To the right of the file list is the content of the file 'transpose.sas'. The content is a SAS script that starts with a comment: '* ID variable is a CHARACTER variable (TEXT);'. The script then defines a data set 'test' and uses the PROC TRANSPOSE statement to create a new data set 'new' from the 'test' data set. The script ends with a RUN statement.

EXAMPLES OF THE SEARCH ENGINE'S USEFULNESS

The search engine is most useful for checking the syntax on certain SAS keywords which have different meanings in different code contexts. Example of these include FORMAT, INPUT & PUT. Also, macro code can be searched for in isolation since the search engine does not suppress special characters such as [% & =].

The user will still have to be 'clever' with their use of the search engine, for example, searching for syntax that is known to be incorrect, to confirm that your own code is correct.

ADVANTAGES

- ✓ Google Search Technology - best in field
- ✓ Character Syntax Search - e.g. [% & =]
- ✓ Fast
- ✓ New ideas

DISADVANTAGES

- ✗ Only Beta – Not 100% bug-free
- ✗ Limited SAS Code Availability – the industry we work in
- ✗ Known Internet Explorer Problems – Firefox & Chrome OK

CONCLUSION

Hopefully this paper will encourage SAS users to trial the Google Labs Code Search engine. Depending on how users currently learn, remember or improve SAS knowledge, this should be tried at least once, and most effectively in conjunction with other resources. Generally the search engine is better for remembering syntax, context of syntax, and macro code rather than learning SAS from scratch. It is advisable to be prepared to be clever with the search criteria, as discussed above.

Google Labs Code Search is still in beta status, so the developers are actively welcoming comments, bugs and suggested improvements in their forum.

Google Labs Code Search, when fully developed and tested, will be one Google product to keep an eye on.

WEBSITE REFERENCES

Google Labs Code Search	http://www.google.com/codesearch
Google Labs	http://www.google.com/labs
Forum	http://groups.google.com/group/Google-Code-Search

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

James Bidwell
Amgen Ltd
1 Uxbridge Business Park
Uxbridge / UB8 1DH
+44 (0) 1895 525451
[bidwell\[at\]amgen\[dot\]com](mailto:bidwell[at]amgen[dot]com)

Brand and product names are trademarks of their respective companies.