
buyCheap (Craigslist)

MBA 614 Final Project

Christopher R. Colton

Executive Summary

One of my hobbies is searching for “stuff” on Craigslist. Like many people I noticed that I search for primarily the same stuff looking to see if anything new has been added. I also found that I do this on a pretty regular basis (sometimes multiple times a night). “buyCheap” is an excel workbook with a vba program that aims to automate that process.

Specifically buyCheap was designed to fulfill the following requirements:

- Allow the user to search Craigslist for anything in housing & for sale including sub categories (including min/max price, has image, No. of bedrooms)
- Allow the user to select how frequently buyCheap will search for results
- buyCheap will send an email to the user when new “stuff” is added
- The program would only run while excel was open

In completing this project I initially divided the program into a few basic functions: user interface, search, email, and updating. In general the concept was easy to think about, reasonable to design, hard to implement, and difficult to get to work in the way intended.

I plan on using buyCheap primarily to search for cars and apartments. (The search for cars will be for a specific car likely at a wide range of price updating daily. On the other hand in looking at houses it will be a very narrow search updating multiple times during the day). The program could be used or modified in a number of ways to fulfill specific user need when searching Craigslist.

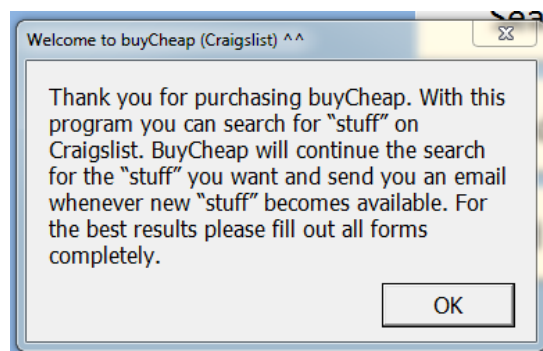
Key words: Craigslist, gmail, search, automatic updating, nested loop, awesome

Walkthrough and implementation documentation

A walkthrough has been included to facilitate discussion of implementation documentation.

What: Welcome message

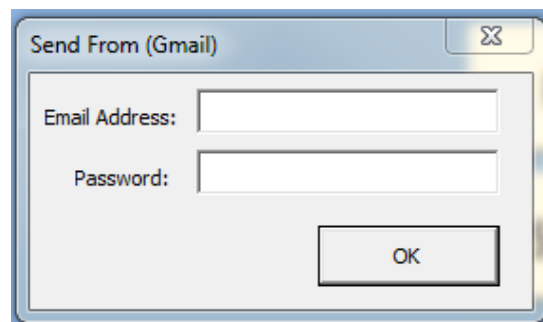
How: Userform formatted as a message box, opens when the workbook is initiated. “OK” button unloads the Userform.



What: The user designates an email account which will be used to send the search results.

How: Similar to what we did in class.

Changes made to the sub procedure will be discussed in detail later.

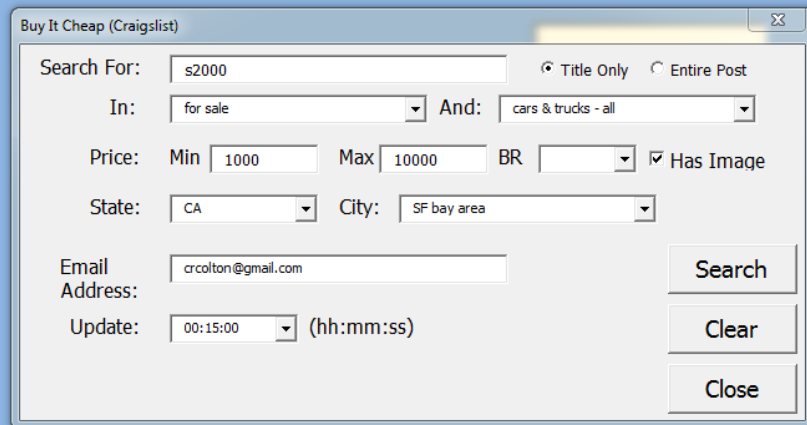


For personal use I disabled this feature and hard coded an email account I created to send the search results. "Ok" button hides the userform.

What: Search userform

Buy it Cheap (Craigslist)

Search For: s2000
Title only: X Entire post:
In: for sale
And: cars & trucks - all
Price
Min: 1000
Max: 10000
BR:
Has image: X
State: CA
City: SF bay area
Email: rcolton@gmail.com
Update: 00:15:00

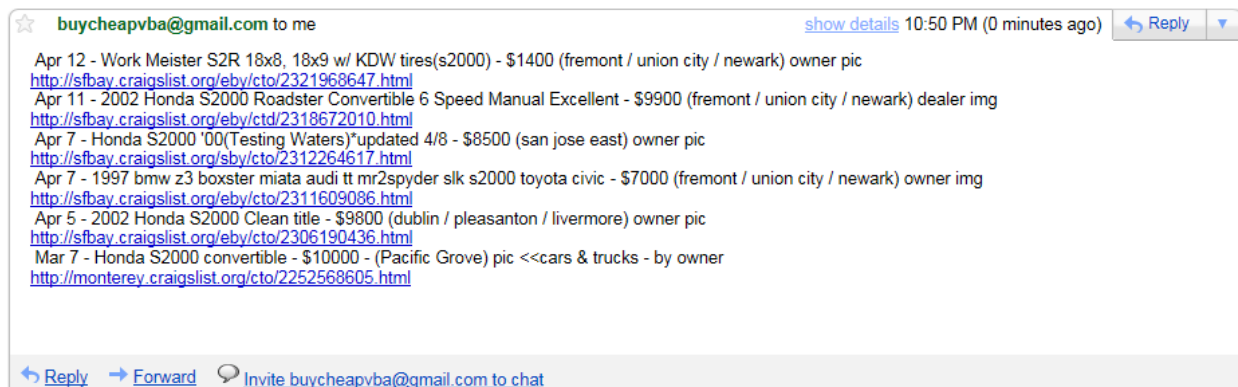


How: I designed a userform to collect the data required to run any type of search in for sale and housing including sub categories on Craigslist. When the user enters data it is written to the excel worksheet. Additional fields were added to allow the user to specify an email address to have the search results sent to and how frequently they would like buyCheap to access Craigslist and import search results. Given that each of the command buttons runs a procedure (or a number of procedures) I will discuss them individually.

What: Search procedures

How: Using the data from the search userform I recreate the url needed to access the same search results on Craigslist. This was done by matching each field with the related piece of the url code. For example if you search for something in SF bay area the url would be <http://sfbay.Craigslist.org/>. Each of the other search criteria builds on after that. Then using the web agent from class I access the sight and pull in all of the search results. These results are stored on the worksheet "searchResults", after which they are paired with the corresponding html links needed to access each individual posting. The data is then cleaned and sorted. The first time buyCheap runs a search it will send you all of the links pulled (up to the first 100 visible on Craigslist).

Search Results For: s2000 Inbox | X



buycheapvba@gmail.com to me show details 10:50 PM (0 minutes ago) Reply

Apr 12 - Work Meister S2R 18x8, 18x9 w/ KDW tires(s2000) - \$1400 (fremont / union city / newark) owner pic
<http://sfbay.craigslist.org/eby/cto/2321968647.html>
Apr 11 - 2002 Honda S2000 Roadster Convertible 6 Speed Manual Excellent - \$9900 (fremont / union city / newark) dealer img
<http://sfbay.craigslist.org/eby/cto/2318672010.html>
Apr 7 - Honda S2000 '00(Testing Waters)*updated 4/8 - \$8500 (san jose east) owner pic
<http://sfbay.craigslist.org/sby/cto/2312264617.html>
Apr 7 - 1997 bmw z3 boxster miata audi tt mr2spyder slk s2000 toyota civic - \$7000 (fremont / union city / newark) owner img
<http://sfbay.craigslist.org/eby/cto/2311609086.html>
Apr 5 - 2002 Honda S2000 Clean title - \$9800 (dublin / pleasanton / livermore) owner pic
<http://sfbay.craigslist.org/eby/cto/2306190436.html>
Mar 7 - Honda S2000 convertible - \$10000 - (Pacific Grove) pic <<cars & trucks - by owner
<http://monterey.craigslist.org/cto/2252568605.html>


Reply Forward Invite buycheapvba@gmail.com to chat

I was unable to find a way to send the hyperlink attached to the description so instead it sends the description followed by the appropriate link. As part of the search function buyCheap compares search results from each update.

It does this by first copying the initial search results to the worksheet "searchResults(old)". Each subsequent search pulls data into "searchResults" and checks for duplicates. These duplicates are deleted from "searchResults" and anything remaining is a new result which is then sent to the specified email address. For the next search anything that was new is added to "searchResults(old)". "searchResults(old)" is designed to hold all matching results, the new results are added to the first empty row in the worksheet.

Given that new data is only available if a Craigslist user makes a new post I have altered the maximum price on our original search which shows that the sort procedure inside the search procedure is in fact functioning correctly.

Search Results For: s2000 Inbox | X

 **buycheapvba@gmail.com** to me show details 10:50 PM (38 minutes ago) Reply

Apr 12 - Work Meister S2R 18x8, 18x9 w/ KDW tires(s2000) - \$1400 (fremont / union city / newark) owner pic
<http://sfbay.craigslist.org/eby/cto/2321968647.html>

Apr 11 - 2002 Honda S2000 Roadster Convertible 6 Speed Manual Excellent - \$9900 (fremont / union city / newark) dealer img
<http://sfbay.craigslist.org/eby/cto/2318672010.html>


Apr 7 - Honda S2000 '00(Testing Waters)*updated 4/8 - \$8500 (san jose east) owner pic
<http://sfbay.craigslist.org/sby/cto/2312264617.html>

Apr 7 - 1997 bmw z3 boxster miata audi tt mr2spyder slk s2000 toyota civic - \$7000 (fremont / union city / newark) owner img
<http://sfbay.craigslist.org/eby/cto/2311609086.html>

Apr 5 - 2002 Honda S2000 Clean title - \$9800 (dublin / pleasanton / livermore) owner pic
<http://sfbay.craigslist.org/eby/cto/2306190436.html>

Mar 7 - Honda S2000 convertible - \$10000 - (Pacific Grove) pic <<cars & trucks - by owner
<http://monterey.craigslist.org/cto/2252568605.html>

Reply Forward Invite buycheapvba@gmail.com to chat

 **buycheapvba@gmail.com** to me show details 11:29 PM (0 minutes ago) Reply

Apr 12 - clean title 2001 honda s2000 - \$13500 (san jose north) owner pic
<http://sfbay.craigslist.org/sby/cto/2321111172.html>

Apr 12 - HONDA S2000 2004 - \$11500 (south san francisco) owner pic
<http://sfbay.craigslist.org/pen/cto/2320861089.html>

Apr 11 - 2003 Honda S2000 Base - \$10530 (san jose east) owner pic
<http://sfbay.craigslist.org/sby/cto/2318853558.html>

Apr 9 - 2002 Honda S2000, stock, 67k, adult owned - \$12500 (santa rosa) owner pic
<http://sfbay.craigslist.org/nby/cto/2316135856.html>

Apr 9 - 2002 Berlina Black S2000 - \$12000 (fremont / union city / newark) owner pic
<http://sfbay.craigslist.org/eby/cto/2314370519.html>

Apr 6 - Hellaflush S2000 Ap1 - \$12000 (san leandro) owner pic
<http://sfbay.craigslist.org/eby/cto/2309760529.html>

Apr 5 - 2003 Honda S2000 spa yellow, Clean Title, Moderate mods, F22C Motor - \$13500 (san mateo) owner pic
<http://sfbay.craigslist.org/pen/cto/2307829642.html>

From the screen capture below you can see that the buyCheap is getting new data every 15 minutes and sending new results to the user. Normally if there are no new results then no email will be sent, this is just given as evidence that the program updates at the user specified time.

Search Results For: s2000 Inbox | X

☆ buycheapvba@gmail.com	Apr 12 - Work Meister S2R 18x8, 18x9 w/ KDW tires(s2000) - \$1400 (fremont / u...	10:50 PM (30 minutes ago)
☆ buycheapvba@gmail.com		11:05 PM (15 minutes ago)
☆ buycheapvba@gmail.com to me	show details	11:20 PM (0 minutes ago) Reply

[Reply](#) [Forward](#) [Invite buycheapvba@gmail.com to chat](#)

What: Clear

How: The clear button resets the search userform. To do this a procedure deletes data saved on the worksheet "New Search" and unloads and re-initializes the search userform. It also resets the variable which stores the code to run automatic updates.

What: Close

How: Although it seems intuitive that close would shut down all the vba code running, this is not the case for this program. Due to the fact that this program is designed to search Craigslist at regular intervals I decided it would be better to hide the search userform than to unload it. Again if a user wanted to reset their search then they would use the clear button and not the close.

What: Search / Read Me / Send From

Buy it Cheap (Craigslist)

Search For:		<div>Search</div> <div>Read Me</div> <div>Send From</div>
Title only:	Entire post: X	
In:		
And:		
Price		
Min:		
Max:		
BR:		
Has image:		
State:		
City:		
Email:		
Update:		

How: After the user closes or hides the search userform they will see the worksheet which holds the search criteria as well as three buttons: search, read me, and send from. Search pulls back up the search userform. Read me opens a userform with some helpful hints on using buyCheap. Lastly send from reopens the userform that allows the user to specify a gmail account that will send the search results.

Learning and conceptual difficulties

In general the concept was easy to think about, reasonable to design, hard to implement, and difficult to get to work in the way I intended. However, I can honestly say that I am proud of the workbook/program I created for my project. I don't know how many hours I spent working on the project but it seems like the most intensive project I've worked on in the MBA program (I'm not sure if

that's a good thing or not). In completing the project I ran into a number of obstacles and had to get help from other students and the professor in order to resolve them. That being said I believe I was able to learn a lot in the process. To preserve brevity I have chosen to highlight only a few of these obstacles and some of the things I learned to combat them.

What: Cascading combo box

Why: In the search userform I needed to link "In" and "And" together as well as "State" and "City". To do this I load the first combo box when the userform initializes and clear then load the second based on changes to the first box.

What: Loops, Loops, and more Loops

Why: In order to clean, sort, and update search results I used a number of "For" loops to remove blank cells, unimportant data, and to write results to new worksheets. One problem that I encountered in deleting blank rows was that if I was running the loop for $r = 1$ to something and two or more blank rows existed concurrently the loop would delete the first row, the second row would become the new first row and the loop would increment leaving a blank row and the program moving on to check if the next row was empty or not. Although it was not perfect I used a reverse loop (i.e. for $x = 300$ to 1 step -1) to delete empty rows resulting in a list of data that was formatted in the desired way. Another thing I tried to do was to write to the last available cell in a column, similarly I used a reverse loop to find the last cell with data in it and then offset by one and write to that cell.

What: Html code and Links

Why: One of the main goals of this program was to send the user an email with links to their search results. I encountered a number of issues in dealing with search descriptions and html links to the actual page displaying the Craigslist post. The first thing was getting url links from Craigslist. Although the process was similar to what we had done in class it was still very challenging for me to do and I required assistance in order to write code for this part of the project. The second challenge was what to do with the links, my first idea was to copy links and descriptions to different columns in a worksheet and then link them based on location. This would result in the perfect linkage of what was in column 1 with what was in column 2. I decided against doing it this way because of the off chance that something strange happened and the data in one column ended up in the wrong cell linked and then was linked to the wrong posting. I discussed this with some other students and decided on linking descriptions and code together from their position in the html code for a search on Craigslist. This works really good with one small exception, when the web query wizard imports a page some characters in individual descriptions change. For example "—" (double hyphen) imports as a question mark and you can't just replace all the question marks with double hyphens. After a number of different tests resulting in failure I decided to check and see how often people use some of the stranger characters in postings, it turns out not very often. The result is that buyCheap does not link 100% of the search results it is probably closer to 100% most of the time 95% the rest of the time.

What: Update Delay

Why: A seeming small yet important part of the project was finding a way to run the program on a time delay defined by the user. When I started I felt like this was out of scope of my knowledge but I knew that other people had included in prior year projects. After talking with the professor I was assured this would be a simple problem and just to get my code to do a single search and then send the results through email. So that's what I did, I completed the code all except for the updating procedure. I then showed my code to some other students and the professor and found out that I could not do the

desired delayed updating with most of the code residing in the userform itself. The result was that I needed to move large sections of code into the general module, and trouble shoot for another two or three hours to get the code to function properly again. Success, the workbook now did everything I needed it to do. I set it to send me emails every three minutes. I started getting emails every two minutes, every one minute, and then the workbook froze. I guess I should have saved my celebration. The problem was that even though I cleared search criteria including the update interval, I did not actually clear the update command logged in memory. To overcome this inserted a few lines of code in the update cycle that reset the update command in memory if the user presses the clear button or the search button for a second time. I coded this as follows:

```
On Error Resume Next
Application.OnTime timeCheck, "runsearch", , False
On Error GoTo 0

If Range("updatetime2").Value > "" Then
    timeCheck = Now + TimeValue(Range("updatetime2"))
    Application.OnTime timeCheck, "runsearch"
End If
```

Conclusion

I learned a lot and believe that I accomplished the key functionality intended in this project. I plan on using buyCheap primarily to search for cars and apartments. (The search for cars will be for a specific car likely at a wide range of price updating daily. On the other hand in looking at houses it will be a very narrow search updating multiple times during the day). The program could be used or modified in a number of ways to fulfill specific user need when searching Craigslist.