

Shopgoodwill.com Web Scraping Project

Table of Contents

How to Compile and Run Code	2
Main GUI.....	2
File Menu Features.....	3
Input	3
View Inventory	4
Reconstruct Log.....	4
Quit.....	4
Search Menu Features.....	5
Online Search	5
Offline Search.....	5
Edit Menu Features	6
Insert.....	6
Modify	6
Delete	7
HashMap implementation.....	7
Logging.....	7
Output	8

Note: All GUI screenshots are from a Mac computer, but a similar GUI should be displayed for a Windows computer.

How to Compile and Run Code

1. go to the path/directory of where your java files are

2. To compile, type in the command line:

javac *.java

3. To run the project, type in the command line:

java P3Main -i items.txt -o output.txt

(I/O files should be in the same directory as project)

in other words, **java P3Main -i <input file name> -o <output file name>**

The output file name is optional meaning `java P3Main -i input.txt` would also be a valid way of executing the code. If no output file name is provided, then `output.txt` will be the default file name used for output.

Examples of incorrect execution of Java code and ways of handling incorrect execution of Java code:

`java P3Main -o output.txt`, `java P3Main -i -o output.txt`, `java P3Main —>` an error will be displayed on the console telling the user the proper way to run the code

`java P3Main—>` an error will be displayed on the console telling the user the proper way to run the code

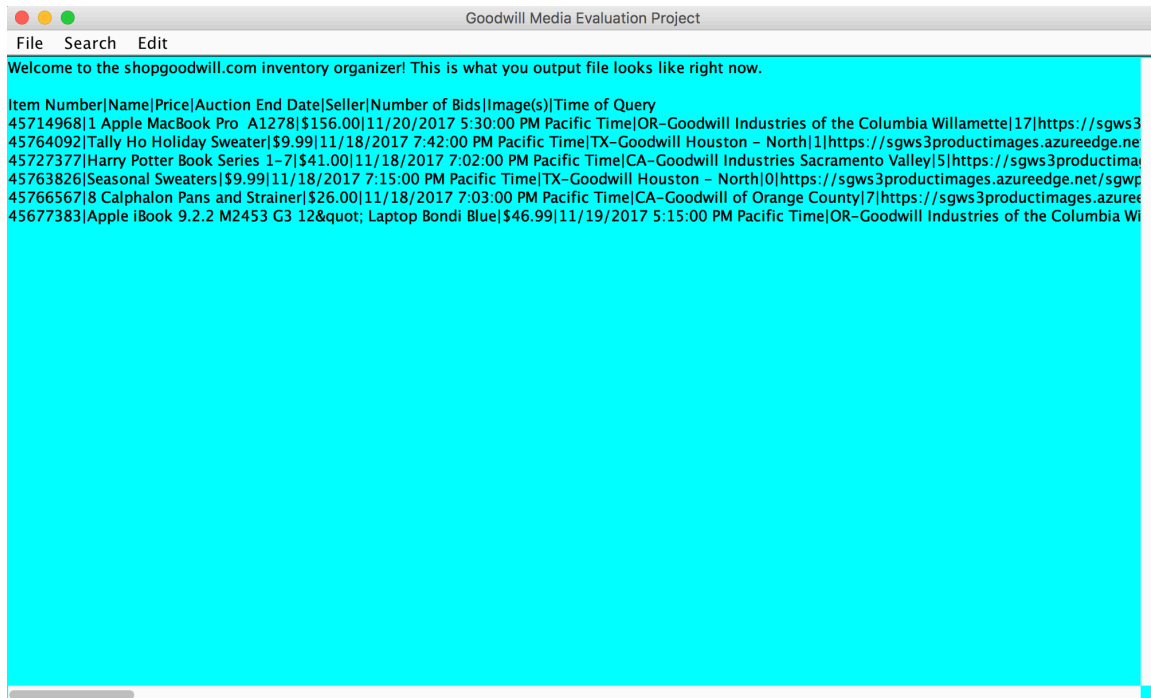
`java P3Main -i —>` an error will be displayed on the console telling the user the proper way to run the code

4. Upon successful compilation/running of code, you should see the MainGUI pop up.

5. Also, if you are performing online searches, there are two hard coded local file paths for the category and seller parameter file in `SearchMenuHandler`. Replace my hard coded local file paths with your path.

Main GUI

When first running the program, the MainGUI will display all of the attributes obtained from the item numbers in the input file corresponding to items in `shopgoodwill.com`. These attributes include item number, name, price, auction end date, number of bids, images, and time of query. You will see on the GUI a preview of what your output file is going to look like. That way you can see what is in your inventory and decide if there is anything else you would like to add, modify, or delete.



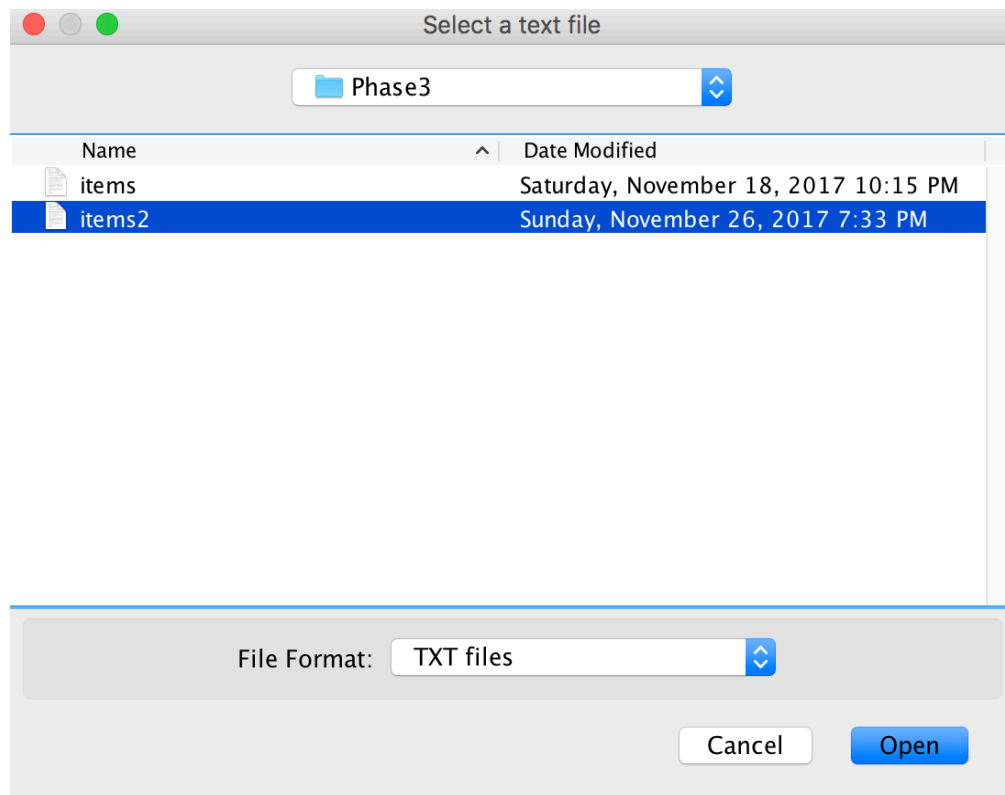
The MainGUI will also display search results of the online search that you performed. Details of an online search are provided in the online search section of this documentation. To go back to your current inventory display, simply go to File -> View Inventory.

The JMenu up top provides many other operations that you can perform while running your code. (Note: if you do not see the JMenu when the GUI starts, drag the bottom right corner of the GUI screen)

File Menu Features

Input

By selecting "Input" under the File menu, a JFileChooser will be displayed and from there, you can select a text file with the proper input format (meaning the input should be a list of item numbers you wish to add to your inventory). Once you select your correctly formatted input file, you will then see those items added to your inventory.



View Inventory

By selecting “View Inventory” under the File menu, the MainGUI will be redirected to the current inventory of the user. This feature is useful after performing an online search.

Reconstruct Log

By selecting “Reconstruct Log” under the File menu, a JFileChooser will be displayed asking for a valid reconstruct log text file. Then the code will read the reconstructed log the same way as adding your original input file to the GUI.

Quit

By selecting “Quit” under the File menu, the program will terminate, and you can obtain your resulting inventory from the output file generated. The reconstruct log file is generated here as well.

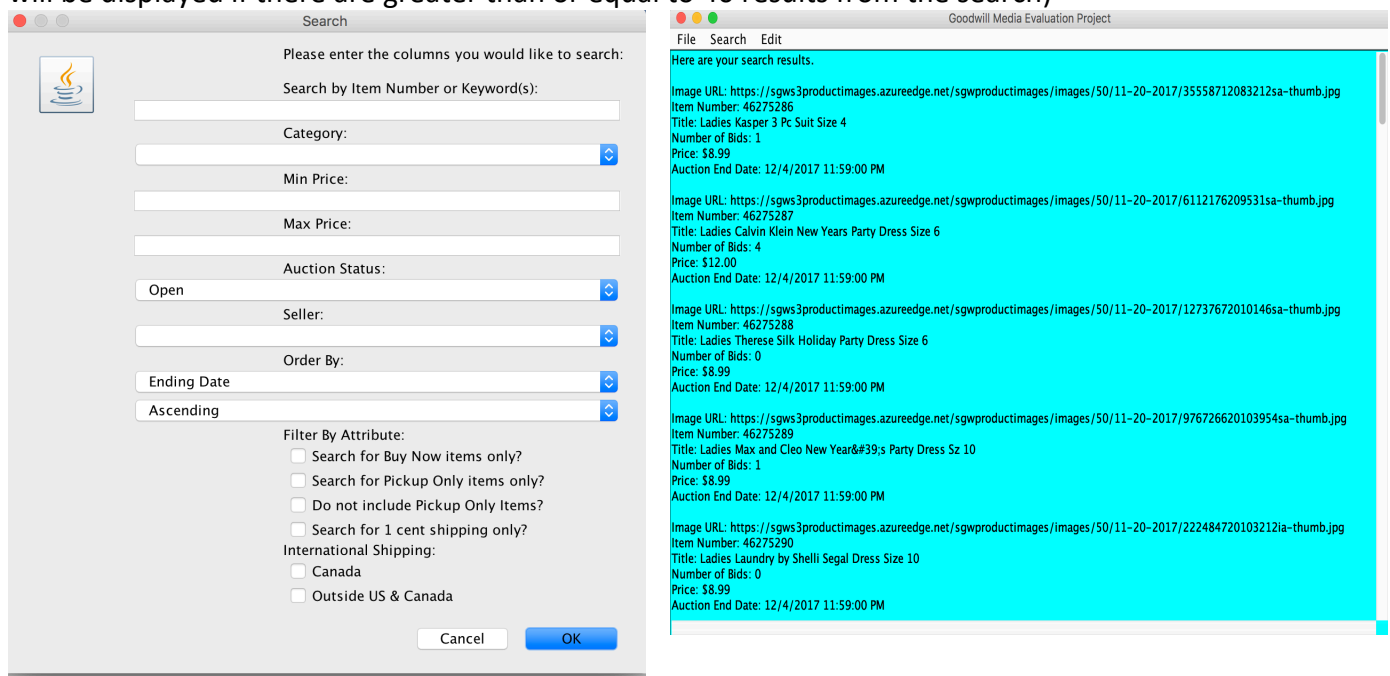
In addition, the images that are stored in the inventory will be saved into your computer. This is done by processing each image URL for each inventory item and saving the image into your computer. The files should be saved in the same folder as the folder you ran your code on. The image name will correspond to the image that is in the URL. For example, the image URL,

<https://sgws3productimages.azureedge.net/sgwproductimages/images/146/12-5-2017/289348756272789.JPG> will produce a JPG file named 289348756272789.JPG.

Search Menu Features

Online Search

To perform an online search, on the MainGUI JMenu, select Search -> Online Search. You should see the following search GUI below. The search parameters are similar to the search parameters provided in the Advanced Search of shopgoodwill.com. Once you fill in your search filters, the results will be displayed on the MainGUI like the one below. (only the first 40 results will be displayed if there are greater than or equal to 40 results from the search)

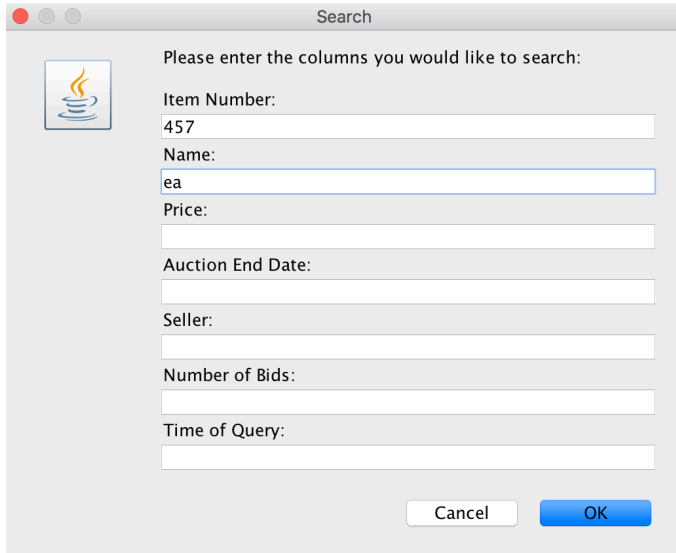


Offline Search

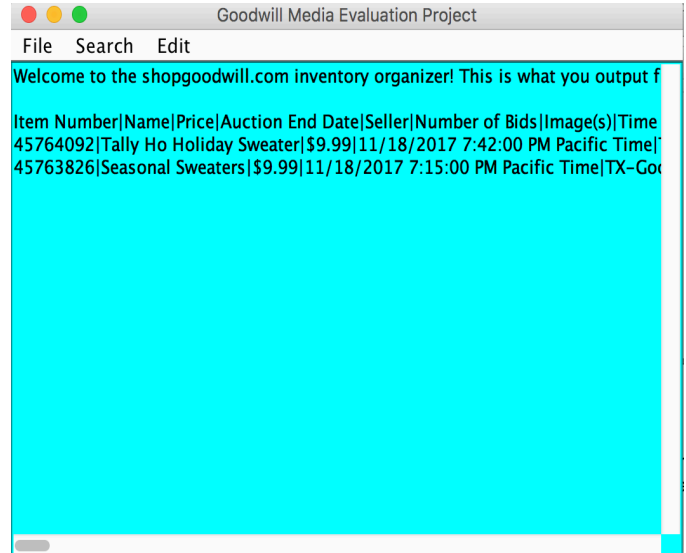
The offline search will filter through each item in your inventory based on item number, name, price, auction end date, seller, number of bids, and/or time of query. Note that all entries are case insensitive.

Basically, when the offline search is called to perform an offline search, all of its search parameters will be placed into an array list with assigned indices based on the order of output in SearchMenuHandler. Then in HashMapQuery, the HashMap containing the current inventory is iterated and each row (or data entry) is placed into a candidate TreeMap to see if this entry satisfies the search requirements. If the search requirements are satisfied, then whatever is in the candidate HashMap will be placed in the GUI. Here is an example illustrating how this works

with the item number and name column. From the example below, any entry that contains “457” in its item number and “ea” in its item name is placed into the GUI as a result of the offline search.



A dialog box titled "Search" with a Java logo icon. It contains a label "Please enter the columns you would like to search:" followed by several input fields: "Item Number:" (containing "457"), "Name:" (containing "ea"), "Price:", "Auction End Date:", "Seller:", "Number of Bids:", and "Time of Query:". At the bottom are "Cancel" and "OK" buttons.



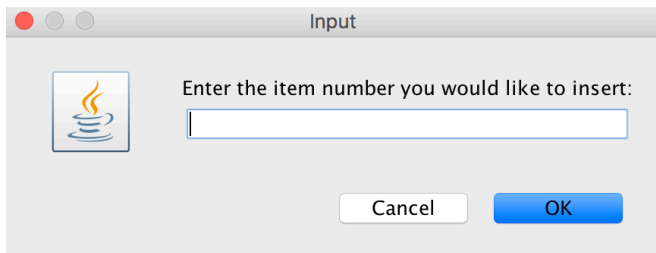
A window titled "Goodwill Media Evaluation Project" with a menu bar (File, Search, Edit). The main area displays a welcome message and a table of search results.

Item Number	Name	Price	Auction End Date	Seller	Number of Bids	Image(s)	Time
45764092	Tally Ho Holiday Sweater	\$9.99	11/18/2017 7:42:00 PM Pacific Time				
45763826	Seasonal Sweaters	\$9.99	11/18/2017 7:15:00 PM Pacific Time	TX-Goo			

Edit Menu Features

Insert

Insertion is simply performed by entering the item number that you would like to insert. Here is an example. The JOOptionPane displayed below is what you will see if you go to Edit -> Insert. Insertion of items will also retrieve data from the corresponding shopgoodwill url like when the code first executes with taking the item numbers in the input and retrieving information from their respective urls.



A dialog box titled "Input" with a Java logo icon. It contains a label "Enter the item number you would like to insert:" followed by a single input field. At the bottom are "Cancel" and "OK" buttons.

Modify

Modify works by editing certain columns in your inventory. Those columns are specified in the GUI used for entering columns you would like to modify. To do this, go to Edit -> Modify

Modify

Please enter the columns you would like to modify: (Item number must be specified)

Item Number:

Name:

Price:

Auction End Date:

Seller:

Number of Bids:

Cancel OK

Delete

Delete will delete an entry in the inventory by item number. Below is the GUI for deletion when you go to Edit -> Delete.

Input

Enter the item number you would like to delete:

Cancel OK

HashMap implementation

The HashMap used for keeping track of inventory is a nested HashMap. The key of the outer HashMap is the item number and the value of the outer hash map is a TreeMap containing the attributes for that item number. The keys of the inner tree map correspond to the ordering of the attributes for each item and the value is the value of that attribute (ex. Key "1" corresponds to the item name/title while the value is the actual name/title)

Logging

Format for log file name:

logMM-DD-YYYY HH_MM_SS.txt where MM-DD-YYYY represents the month, day, and year of when you ran the code and HH_MM_SS represents the time of when this file was generated

(ex. log12-06-2017 13_10_43.txt)

The log file contains all of the operations (such as insert, modify, and delete) performed in the GUI.

The text file will be in the following format:

QUERY OPERATION

Item Number|Name|Price|Auction End Date|Seller|Number of Bids|Image(s)|Time of Query

where QUERY OPERATION = INSERT, MODIFY, or DELETE

A special log file for reconstruction is also created. This reconstruct log file contains a list of item numbers after using the GUI operations.

Format for reconstruct log file name:

logreconstructMM-DD-YYYY HH_MM_SS.txt where MM-DD-YYYY represents the month, day, and year of when you ran the code and HH_MM_SS represents the time of when this file was generated

(ex. logreconstruct12-06-2017 13_19_06.txt)

Output

The output file generated contains the final set of inventory after all query operations and searches have been performed. The delimiters provided for each item allow you to convert this text file into Excel.

Here is a sample

```
Item Number|Name|Price|Auction End Date|Seller|Number of Bids|Image(s)|Time of Query
45714968|1 Apple MacBook Pro A1278|$156.00|11/20/2017 5:30:00 PM Pacific Time|OR-Goodwill
Industries of the Columbia
Willamette|17|https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-13-
2017/5675901393247to.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-
13-
2017/22316701393247to.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-
13-
2017/25217171393247to.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-
13-
2017/70969781393247to.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-
13-
2017/83565451393247to.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-
13-2017/72043391393247to.JPG:|12/03/2017 22:08:48
45764092|Tally Ho Holiday Sweater|$9.99|11/18/2017 7:42:00 PM Pacific Time|TX-Goodwill Houston -
North|1|https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-
2017/8571740141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
14-
2017/5844340141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
14-
2017/1772206141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
14-
2017/3854791141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
14-
2017/722879141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
14-
2017/7701166141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
14-
2017/8501991141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-
```


-14-
2017/5313229141183011.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/132510141183011.jpg:|12/03/2017 22:08:48
45727377|Harry Potter Book Series 1-7|\$41.00|11/18/2017 7:02:00 PM Pacific Time|CA-Goodwill Industries Sacramento Valley|5|https://sgws3productimages.azureedge.net/sgwproductimages/images/162/11-13-2017/935344713141918nL.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/162/11-13-2017/606147313141918nL.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/162/11-13-2017/377095413141918nL.JPG:|12/03/2017 22:08:48
45763826|Seasonal Sweaters|\$9.99|11/18/2017 7:15:00 PM Pacific Time|TX-Goodwill Houston - North|0|https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/4259911141102311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/2184978141102311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/9359180141102311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/166132141102311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/7686685141102311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/5223656141104211.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/1279216141104211.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/8658497141104311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/9494744141104311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/5116494141104311.jpg:https://sgws3productimages.azureedge.net/sgwproductimages/images/185/11-14-2017/225275141104311.jpg:|12/03/2017 22:08:48
45766567|8 Calphalon Pans and Strainer|\$26.00|11/18/2017 7:03:00 PM Pacific Time|CA-Goodwill of Orange County|7|https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/10844051412240as.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/75081591412241as.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/2322891412241as.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/91697321412241as.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/30977571412242as.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/87711121412242as.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/2/11-14-2017/18096401412243as.JPG:|12/03/2017 22:08:48
45677383|Apple iBook 9.2.2 M2453 G3 12" Laptop Bondi Blue|\$46.99|11/19/2017 5:15:00 PM Pacific Time|OR-Goodwill Industries of the Columbia Willamette|12|https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-12-2017/83848191273751rd.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-12-2017/781971273751rd.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-12-2017/4268381273751rd.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-12-2017/15212651273752rd.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-12-2017/21869601273752rd.JPG:https://sgws3productimages.azureedge.net/sgwproductimages/images/8/11-12-2017/66432741273752rd.JPG:|12/03/2017 22:08:48

