

# Savvy Realm Web Extractor 1.0

## User Guide

Written by  
Robert G. Marquez

Document Release Number 3

**Author of This Publication:**

Robert G, Marquez

**Designer and Developer of Savvy Realm Web Extractor Software**

Robert G. Marquez

**Software This Book Covers**

This book provides information about Savvy Realm Web Extractor 1.0.0

**Book Release Date:**

This book, Savvy Realm Web Extractor User Guide, was released on March 7, 2019

**Copyright © 2019 by Robert G. Marquez \ Savvy Realm**

All rights reserved. Savvy Realm grants its registered and non-registered users of Savvy Realm Web Extractor the right to make one printed copy of this book for the sole purpose of working with Savvy Realm Web Extractor themselves. This book may not be reprinted for the purpose of re-selling or exchanging of any goods or funds.

**Document Release History**

First release (# 1) March 7, 2019. Release for this printing is March 8, 2019 (#2)

**Trademarks:**

All products and brand names used in this book are trade names, service marks, trademarks, or registered trademarks of their respective owners. Savvy Realm is not associated with any vendor or product mentioned in this book other than Savvy Realm Web Extractor itself.

**LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY:**

SAVVY REALM HAS USED TO ITS BEST ABILITIES TO PREPARE THIS BOOK. SAVVY REALM MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE COMPLETENESS OR ACCURACY OF THE CONTENTS WITHIN THIS BOOK AND SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. THE LIMITATIONS OF ANY WARRANTY ARE DESCRIBED IN THIS PARAGRAPH. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES REPRESENTATIVES OR OTHER SUCH PERSONS OR WRITTEN MATERIALS SELLING THIS BOOK. THE ACCURACY AND COMPLETENESS OF THE INFORMATION CONTAINED IN THIS BOOK ARE NOT GUARANTEED OR WARRANTED TO PRODUCE ANY PARTICULAR RESULTS. SUGGESTIONS OR ADVICE WITHIN THIS BOOK MAY NOT BE SUITABLE FOR EVERY INDIVIDUAL. NEITHER SAVVY REALM, ITS PERSONNEL, THE AUTHOR OF THIS BOOK, OR THE AUTHOR OF SAVVY REALM WEB EXTRACTOR SHALL BE LIABLE FOR ANY LOSS OF PROFIT OR ANY OTHER COMMERCIAL DAMAGES, INCLUDING BUT NOT LIMITED TO SPECIAL, INCIDENTAL, CONSEQUENTIAL, OR OTHER DAMAGES.

**Savvy Realm Web Extractor Software License Statement**

YOUR USE OF THE SOFTWARE, SAVVY REALM WEB EXTRACTOR, DISTRIBUTED WITH THIS LICENSE IS SUBJECT TO ALL OF THE TERMS AND CONDITIONS OF THIS LICENSE STATEMENT. DO NOT USE THIS SOFTWARE IF YOU DO NOT AGREE TO ALL OF THE TERMS AND CONDITIONS OF THIS LICENSE STATEMENT.

This SOFTWARE is protected by copyright law and international copyright treaty. You are allowed to make copies of this software for the sole purpose of backing up the SOFTWARE for your personal use only.

SAVVY REALM MAKES NO REPRESENTATIONS ABOUT THE SUITABILITY OF THE SOFTWARE OR ABOUT ANY CONTENT OR INFORMATION MADE ACCESSIBLE BY THE SOFTWARE, FOR ANY PURPOSE. THE SOFTWARE IS PROVIDED "AS IS" WITHOUT EXPRESS OR IMPLIED WARRANTIES, INCLUDING WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PUPOSE OR NONINFRINGEMENT. THIS SOFTWARE IS PROVIDED GRATUITOUSLY AND, ACCORDINGLY, SAVVY REALM SHALL NOT BE LIABLE UNDER ANY THEORY OR ANY DAMAGES SUFFERED BY YOU OR ANY USER OF THE SOFTWARE. SAVVY REALM IS UNDER NO OBLIGATION TO ISSUE UPDATES TO THE SOFTWARE.

The term SOFTWARE in this license agreement refers to the software called Savvy Realm Web Extractor designed and developed by Robert G Marquez, business owner of Savvy Realm. Savvy Realm reserves the right at any time to alter prices, features, specifications, capabilities, functions, licensing terms, release dates, general availability or any other element of the commercial release.

Title, ownership rights, and intellectual property rights in and to the SOFTWARE shall remain with Robert G. Marquez \ Savvy Realm and/or its supplies.

You agree to abide by the copyright law and all other applicable laws of the United States including, but not limited to, export control laws. You acknowledge that the SOFTWARE in source code form remains a confidential trade secret of Savvy Realm and therefore you agree not to modify the SOFTWARE or attempt to decipher, decompile, disassemble or reverse engineer the SOFTWARE, except to the extent applicable laws specifically prohibit such restriction.

Duplication, use, or disclosure by the Government is subject to restrictions set forth in subparagraphs (a) through (d) of the Commercial Computer-Restricted Rights clause at FAR 52.227-19 whenever applicable, or in subparagraph (c) (1) (ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013, and in similar clauses in the NASA AR Supplement. Contractor/ manufacturer is Savvy Realm.

You may not export, re-export, or download the SOFTWARE or any underlying information or technology except in full compliance with all United States and other applicable laws and regulations. In particular, but without limitation, none of the SOFTWARE or underlying information or technology may be exported, reexported, or downloaded into (or to a national or resident of) Cuba, Haiti, Iraq, Libya, Yugoslavia, North Korea, Iran, or Syria or to anyone on the US Treasury Department's list of Specially Designated Nationals or the US Commerce Department's Table of Deny Orders. By downloading the Software, you are agreeing to the foregoing and you are

representing and warranting that you are not located in, under control of, or a national or resident of any such country or on any such list.

SAVVY REALM OR ITS SUPPLIERS SHALL NOT BE LIABLE FOR CONSEQUENTIAL, INCIDENTAL, SPECIAL OR INDIRECT DAMAGES OF ANY SORT, WHETHER ARISING IN CONTRACT, TORT, OR OTHERWISE, EVEN IF SAVVY REALM HAS BEEN INFORMED OF THE POSSIBILITY OF SUCH DAMAGES, OR FOR ANY CLAIM BY ANY OTHER PARTY. THIS LIMITATION OF LIABILITY SHALL NOT APPLY TO LIABILITY FOR DEATH OR PERSONAL INJURY TO THE EXTENT APPLICABLE LAW PROHIBITS SUCH LIMITATION. FURTHERMORE, SOME STATES DO NOT PERMIT THE EXCLUSION OR LIMITATION OF INCIDENTAL OR CONSEQUENTIAL DAMAGES, THEREFORE THIS LIMITATION AND EXCLUSION MAY NOT APPLY TO YOU.

HAZARDOUS, DANGEROUS, OR CRITICAL ACTIVITIES. The SOFTWARE is not fault-tolerant and is not manufactured, designed, or intended for resale or use in any way as on-line control equipment in dangerous or life threatening conditions in environments requiring fail-safe performance for areas such as military defense, emergency response (911), nuclear facilities, aircraft navigation, communication systems, air traffic control, direct life support machines, any medical activities required to protect someone, or weapons systems, where failure of the SOFTWARE could lead directly to death, personal injury, or severe physical or environmental damage. Savvy Realm and its suppliers specifically disclaim any express or implied warranty of fitness for hazardous, dangerous, or critical activities.

Trademarks & Copyright 2019 Savvy Realm.

## Contents

Overview.....	9
System Requirements.....	9
Software Installation .....	9
Accessing Savvy Realm Web Extractor After Installation.....	14
Running from The Desktop Shortcut.....	14
Running from the Start Menu .....	14
Running from the Windows Search Box:.....	15
Uninstalling the Savvy Realm Web Extractor Application:.....	16
Using the Application .....	17
HTML Select List Conversions Window .....	18
1.0 Conversion Type .....	19
1.1 Radio Button – Extract Text from Option Values .....	19
1.2 Radio Button – Extract Text from Option Labels.....	21
1.3 Radio Button – Convert text list to HTML Select List.....	22
2.0 The Data to Convert Box.....	22
Available Methods To Enter HTML <select> Statements .....	23
2.1 Load File Button.....	23
2.2 Load Sample Button .....	24
2.3 Paste Button .....	24
2.4 Convert Button .....	24
2.5 Clear Button.....	24
2.6 Help Button.....	24
3.0 The Results Box.....	25
3.1 Copy Button.....	25
3.2 Clear Button.....	25
3.3 Save Results .....	25

## Table Of Figures

Figure 1 Suggested Installation Download Location .....	10
Figure 2 Installation Window (Welcome).....	10
Figure 3 License Agreement .....	11
Figure 4 Installation Window (Features) .....	11
Figure 5 Select installation folder .....	12
Figure 6 Begin Installation .....	12
Figure 7 Installation Window (User Account Control) .....	13
Figure 8 Installation Window (Completion) .....	13
Figure 9 Desktop Shortcut .....	14
Figure 10 Running from the Start Menu .....	14
Figure 11 Running from the Windows Search Box.....	15
Figure 12 Uninstalling the Savvy Realm Web Extractor Application.....	16
Figure 13 Savvy Realm Web Extractor Main Window .....	17
Figure 14 HTML Select List Conversions Window .....	18
Figure 15 Conversion Type Extract Areas.....	19
Figure 16 Conversion Type - Extract text from option values .....	20
Figure 17 Conversion Type - Extract text from option labels.....	21
Figure 18 Conversion Type - Convert text list to html select list.....	22
Figure 19 Load Text File.....	23

## Overview

Savvy Realm Web Extractor (SRWE) is an application built to assist in collecting information from web pages that may be needed for tasks such as reports or databases. In its first release, Savvy Realm Web Extractor can be used to extract data contained in standard formatted HTML select statements such as the following:

```
<SELECT>
<option value = "Mercury">Mercury Planet</ option >
<option value = "Venus">Venus Planet</ option >
<option value = "Earth">Earth Planet</ option >
<option value = "Mars">Mars Planet</ option >
<option value = "Jupiter">Jupiter Planet </ option >
</SELECT>
```

The internet is filled with pages containing information such as lists that are needed for a variety of reasons. Of course, there are many pages where the information displayed on a web page can simply be copied and pasted. But in the case of a drop-down list of values a copy and paste operation won't usually work. This is where the Savvy Realm Web Extractor tool can help. Savvy Realm Web Extractor can extract listed information from HTML select statements in the following:

- 1) Drag Drop a web page URL to SRWE.
- 2) Copy and paste an accessible <select> statement from exposed HTML source or from a published location such as GITHUB.
- 3) Load a file into SRWE which contains <select> statements you want list data extracted from

More details about how a user can use SRWE to perform data list extractions are covered later in this user guide.

## System Requirements

Savvy Realm Web Extractor's minimum run time requirements are as follows:

- Windows 10 (20 megs disc space, 4 megs ram)
- Windows 7 (20 megs disc space, 4 megs ram)

## Software Installation

It is recommended you download the software to your PC into the download folder as shown below:

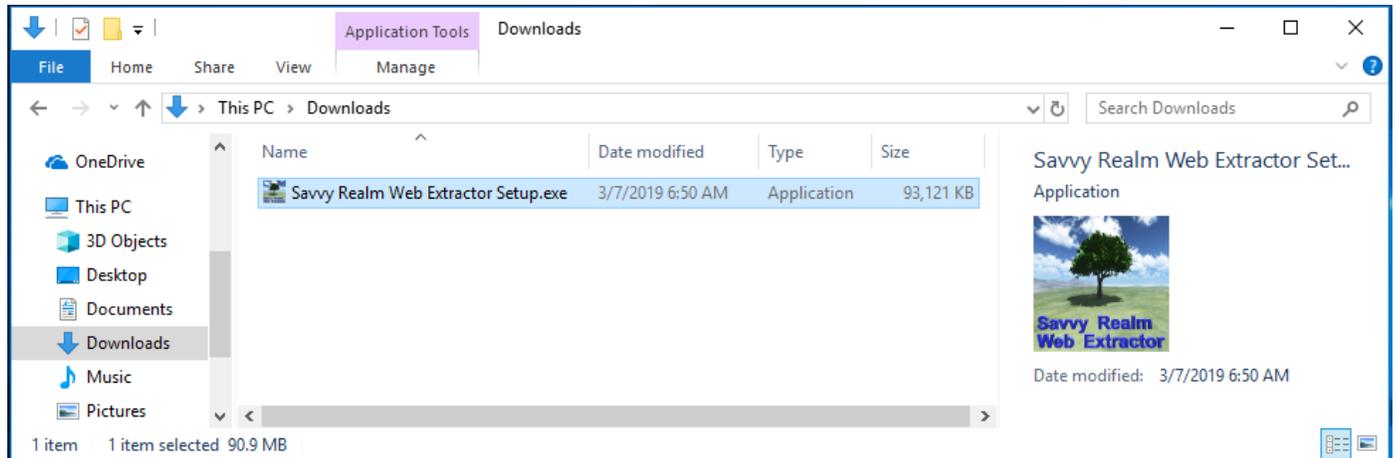


Figure 1 Suggested Installation Download Location

Start the installation by right clicking it and selecting the “Run as Administrator” option. The following window will be shown when the installation starts:

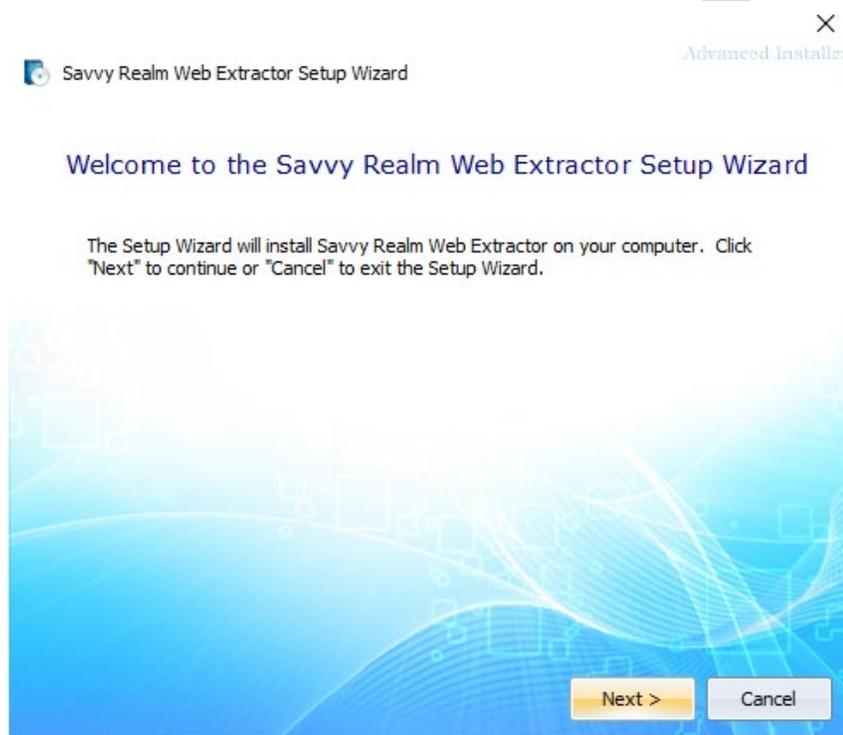


Figure 2 Installation Window (Welcome)

Click the “Next” button to continue.

The following window is the License Agreement. To review it you may print or scroll the window shown here. Click the Accept button if you are in agreement to proceed to the next step.

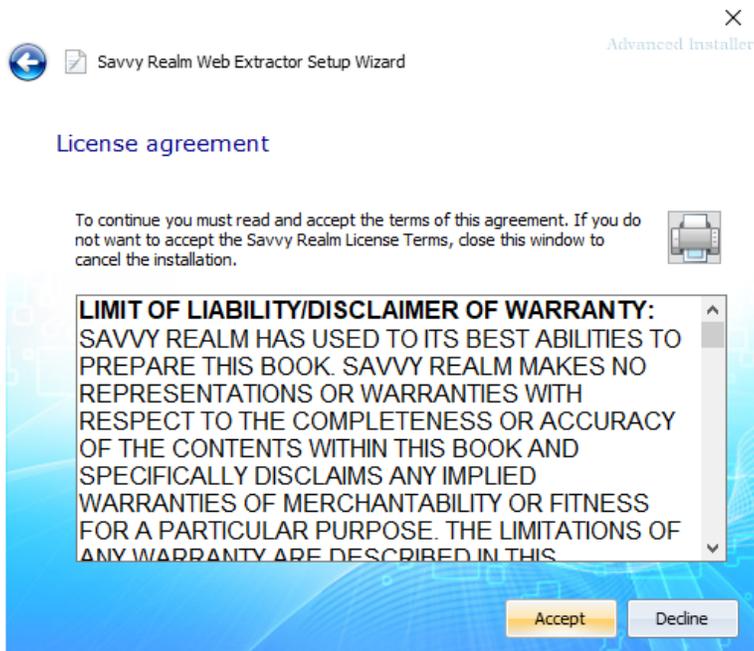


Figure 3 License Agreement

When the following window is shown, click the “Typical” option if you do not need any customizations.

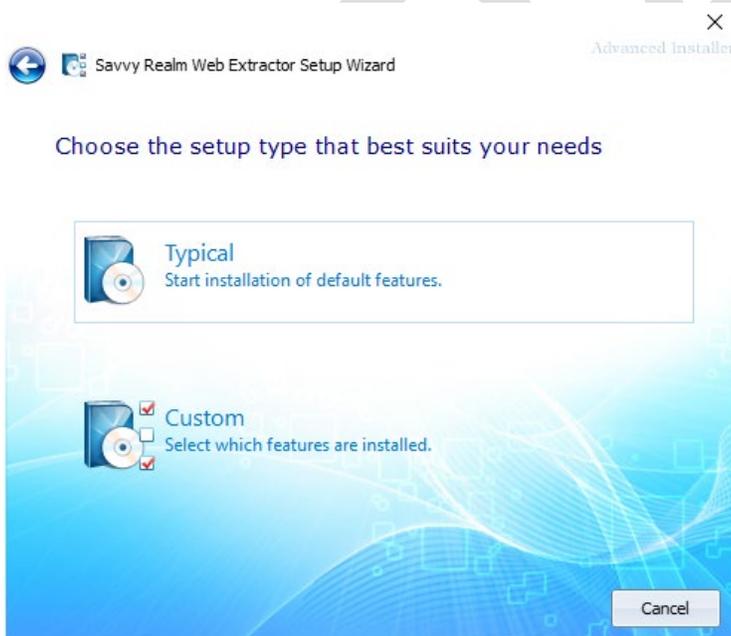


Figure 4 Installation Window (Features)

After clicking the “Typical” option, the window below will be shown. You can change the folder where the application will be installed by clicking the Browse button. To proceed with the default location just Click the “Next” button to continue.

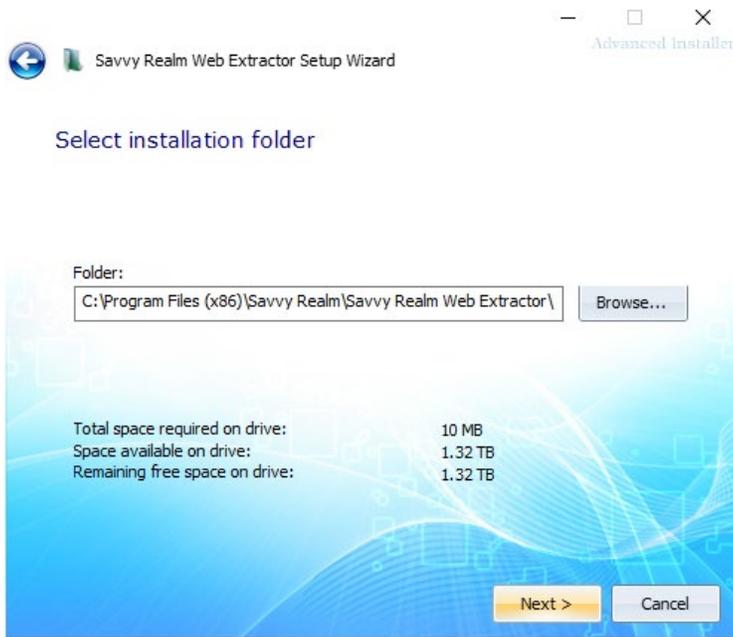


Figure 5 Select installation folder

Click the “Install” button to continue when you see the below window.

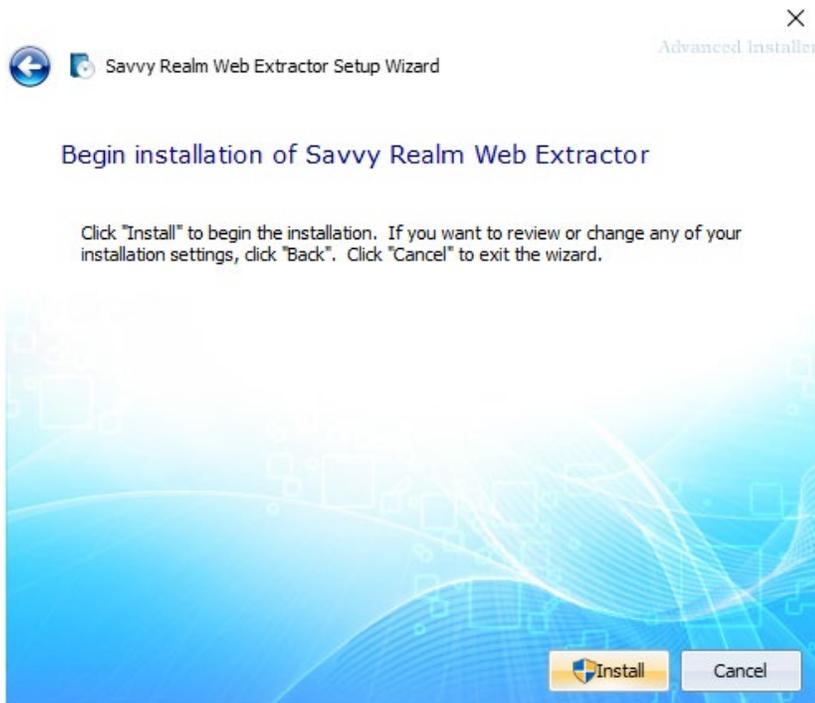


Figure 6 Begin Installation

If the following window is shown, click the “Yes” button to continue.

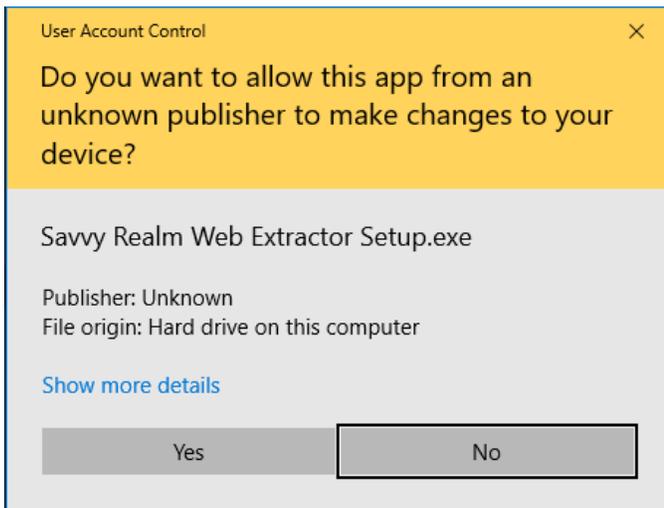


Figure 7 Installation Window (User Account Control)

The installation has been completed when you see the below window. Click the “Finish” button to end the installation now that it has been completed.

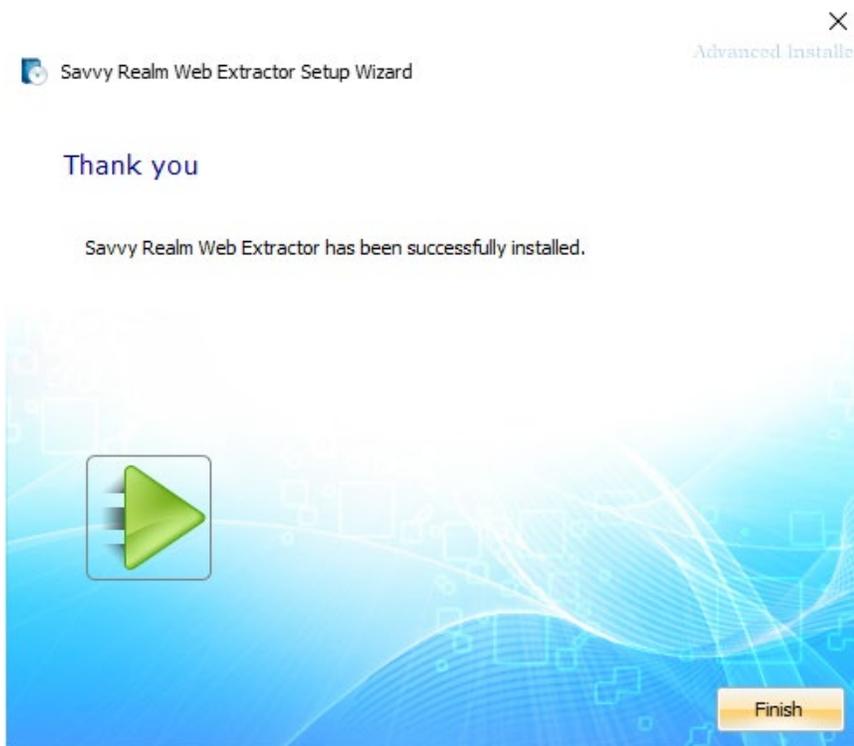


Figure 8 Installation Window (Completion)

## Accessing Savvy Realm Web Extractor After Installation

After the installation has completed you can run the software from a few locations such as the following:

- Desktop
- Start Menu

### Running from The Desktop Shortcut

If a shortcut was created on the desktop, you will find it there as shown below:



Figure 9 Desktop Shortcut

You can launch the application by double clicking this desktop shortcut or right clicking it and selecting “open”.

### Running from the Start Menu

You may run the SRWE application from the Start Menu by following instructions in the below picture for steps 1 through 3:

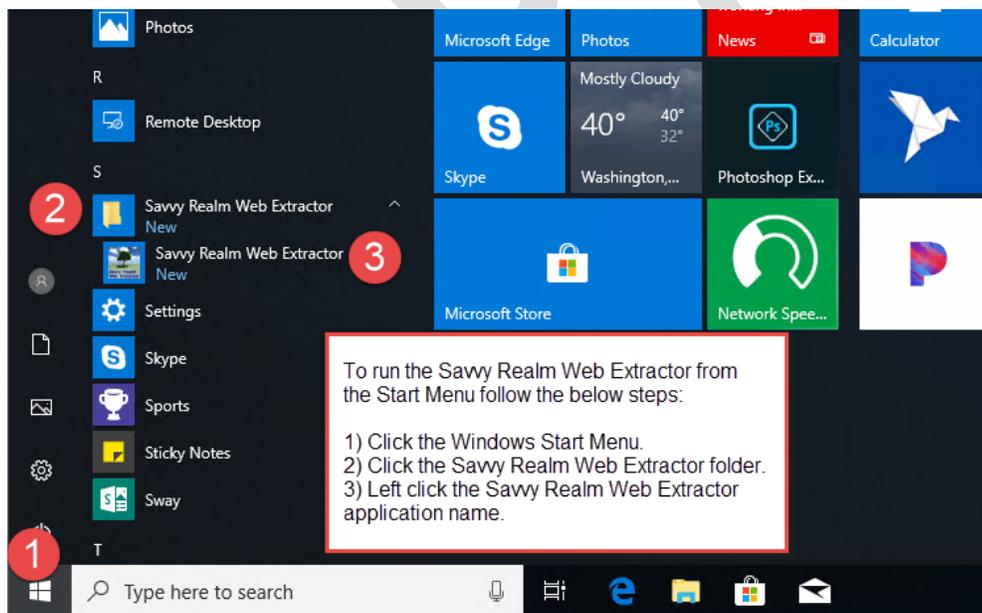


Figure 10 Running from the Start Menu

## Running from the Windows Search Box:

To run the SRWE Application from the Windows Search Box follow the steps shown below:

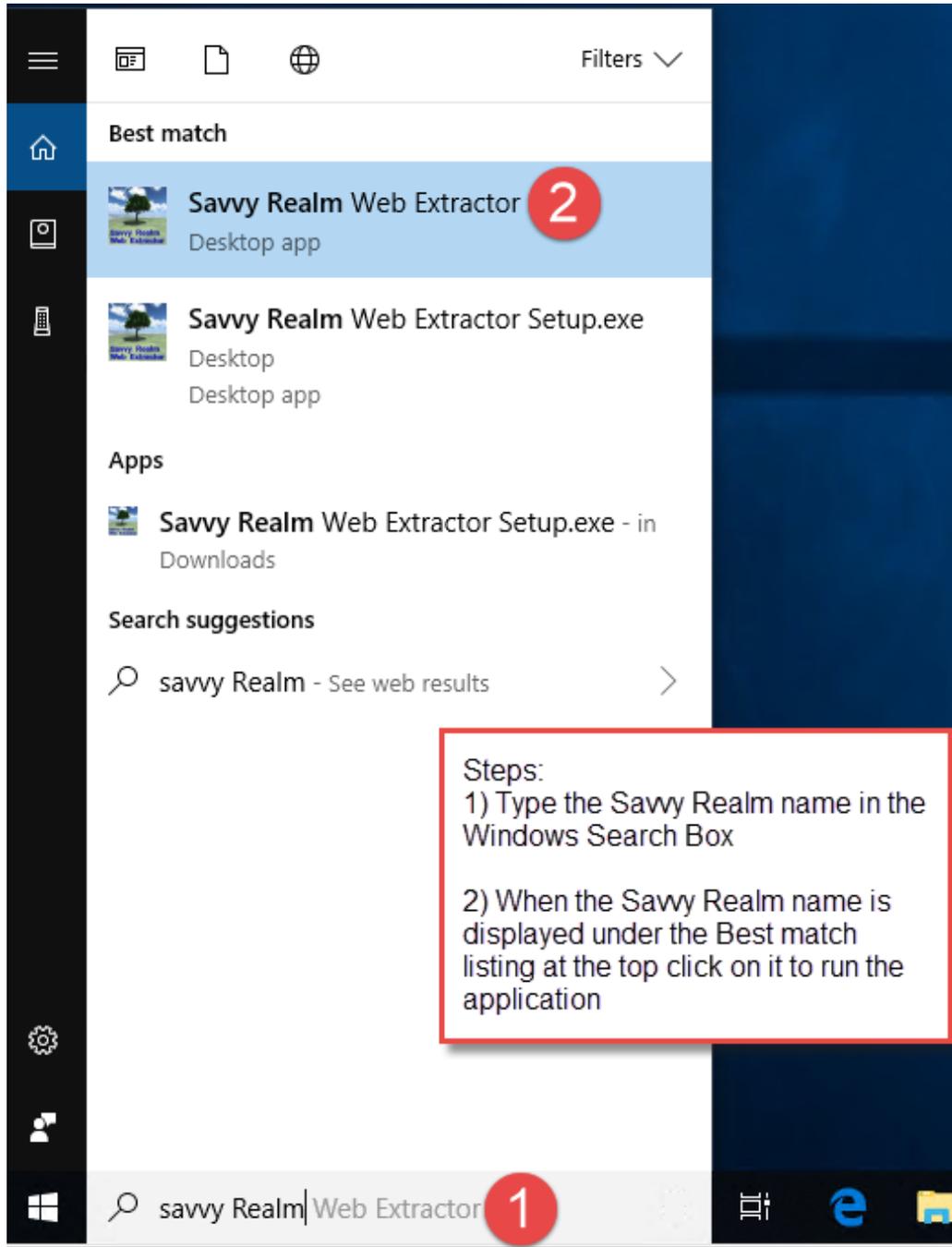


Figure 11 Running from the Windows Search Box

## Uninstalling the Savvy Realm Web Extractor Application:

To uninstall the Savvy Realm Web Extractor, follow the steps shown below:

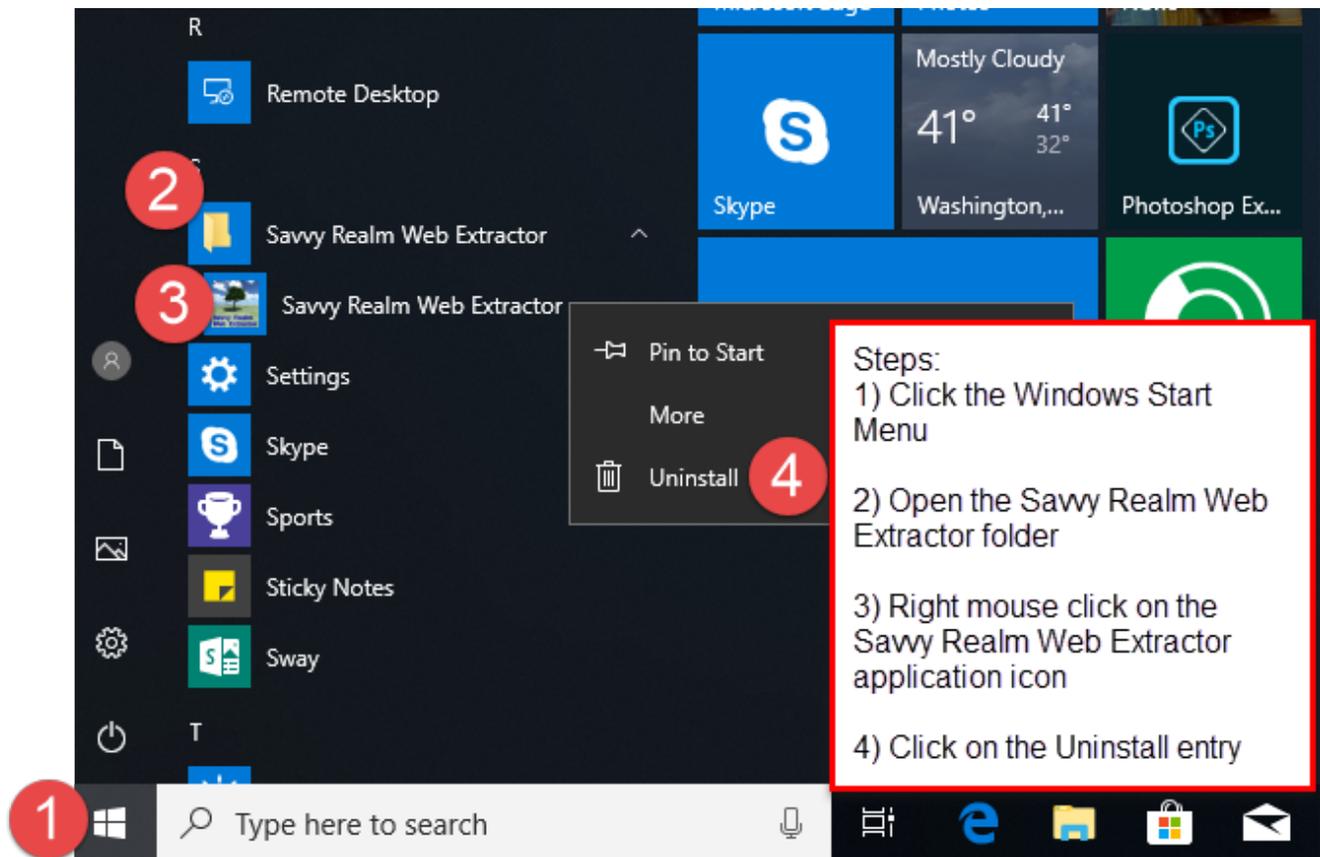


Figure 12 Uninstalling the Savvy Realm Web Extractor Application

## Using the Application

The main window for SRWE will be shown below when you first start the application:

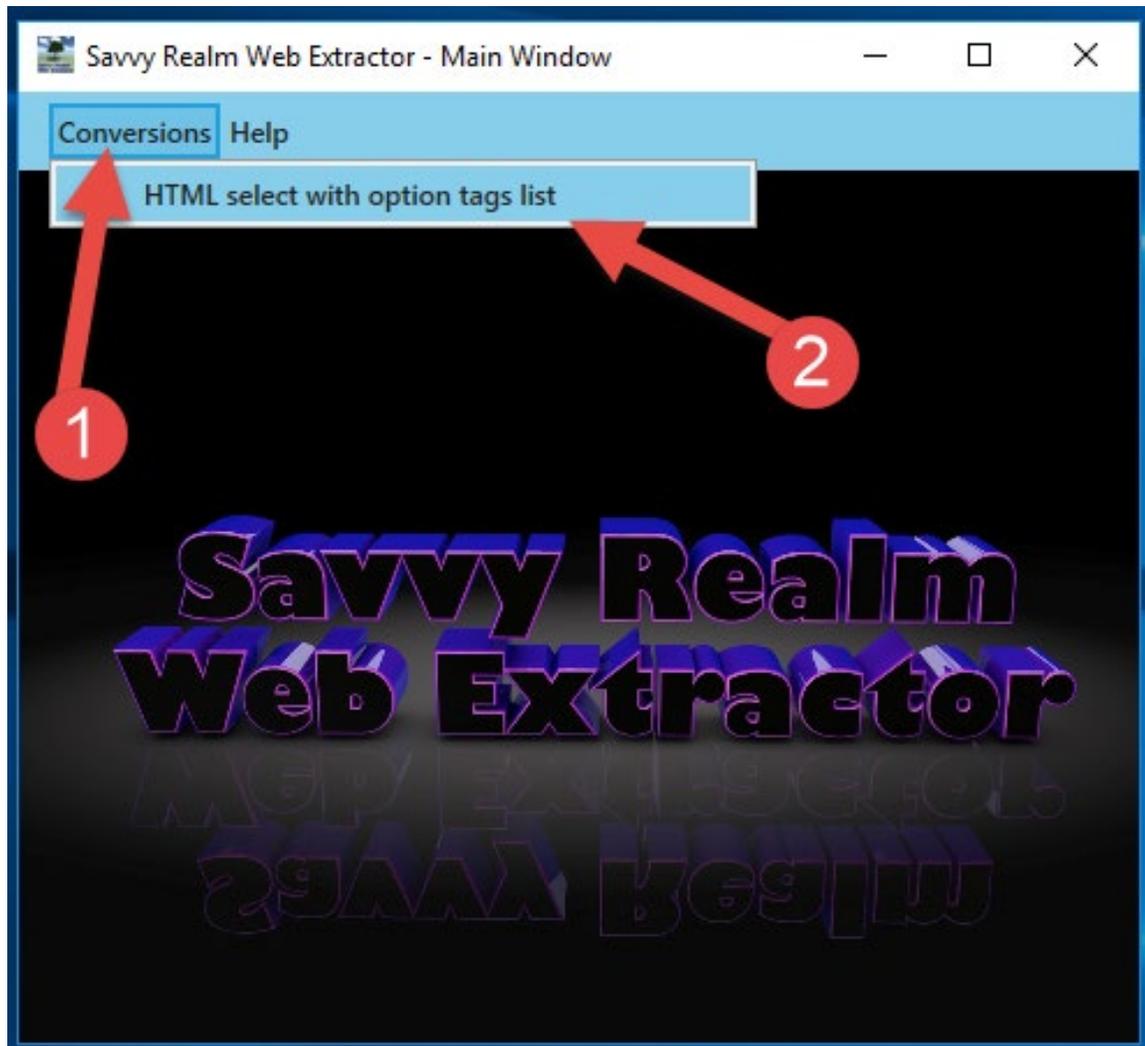


Figure 13 Savvy Realm Web Extractor Main Window

There is only one operation you can perform in this initial version of SRWE and that is performing conversions on HTML <select> lists. To access this part of the application you need to perform the following steps from the main window shown above.

- 1) Click on the Conversions menu item
- 2) After performing step 1, another sub menu will be shown titled “HTML select with option tags list. Click on this menu option.

## HTML Select List Conversions Window

After performing the previous steps, the HTML Select List Conversions window will be shown below:

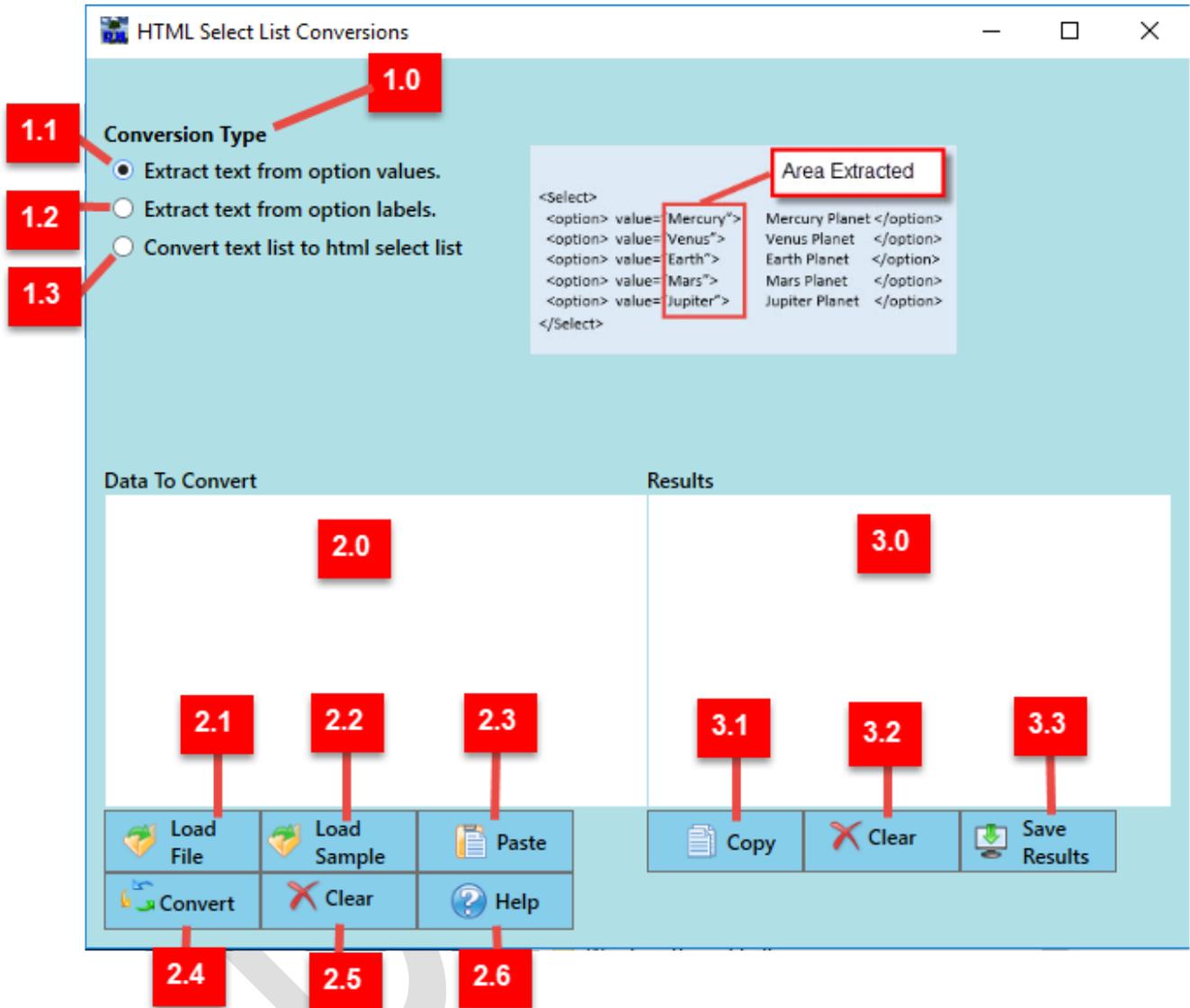


Figure 14 HTML Select List Conversions Window

**Note:** The following pages refer to the above red box labels when describing the objects they point to..

This window is called “HTML Select List Conversions because it is built to convert the values of an HTML <select> statement into a single list based on the <select> values to submit or the <select> values being used to display values to a user. Below is an example of an HTML <select> statement with <option> values:

```
<SELECT>  
<option value = "Mercury"> Mercury Planet </ option >  
<option value = "Venus"> Venus Planet </ option >  
<option value = "Earth"> Earth Planet </ option >  
<option value = "Mars"> Mars Planet </ option >  
<option value = "Jupiter"> Jupiter Planet </ option >  
</SELECT>
```

## 1.0 Conversion Type

The conversion type dictates what data is to be extracted from the HTML Select Option statement. There are two types of data which can be extracted which are the "Option Value" and "Option Label". The Option Value is the value to be submitted for the selected option. The Option Label is the value displayed to a user but is not submitted as the selected value. The below sample HTML shows where these two value types are:

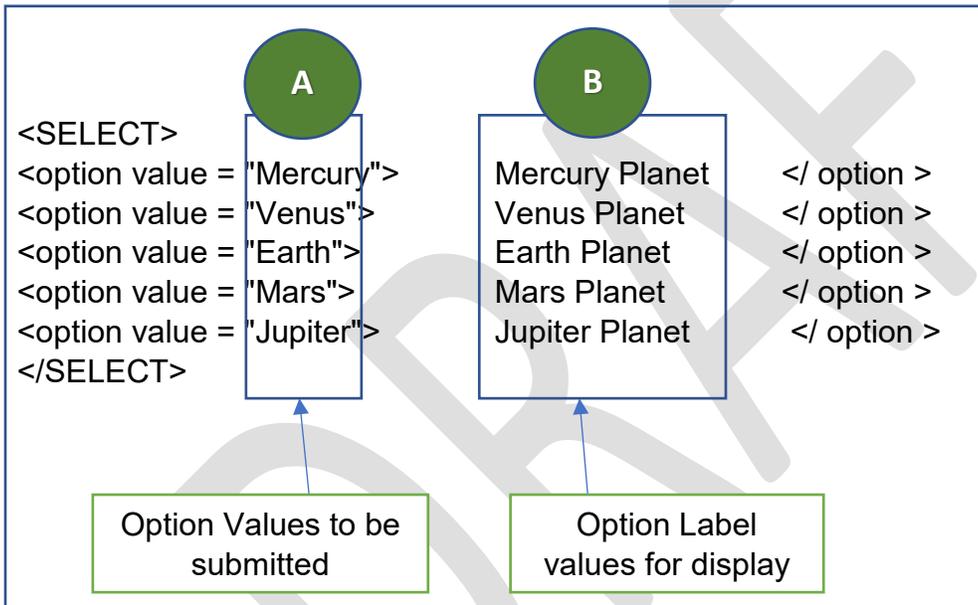


Figure 15 Conversion Type Extract Areas

### 1.1 Radio Button – Extract Text from Option Values

Selecting the radio button Extract text from option values followed by clicking the Convert Button results in values being extracted from the section above marked with the letter "A" which would give you the following list in the Results box:

- Mercury
- Venus
- Earth
- Mars
- Jupiter

The HTML Select List Conversions Window would look similar to the following:

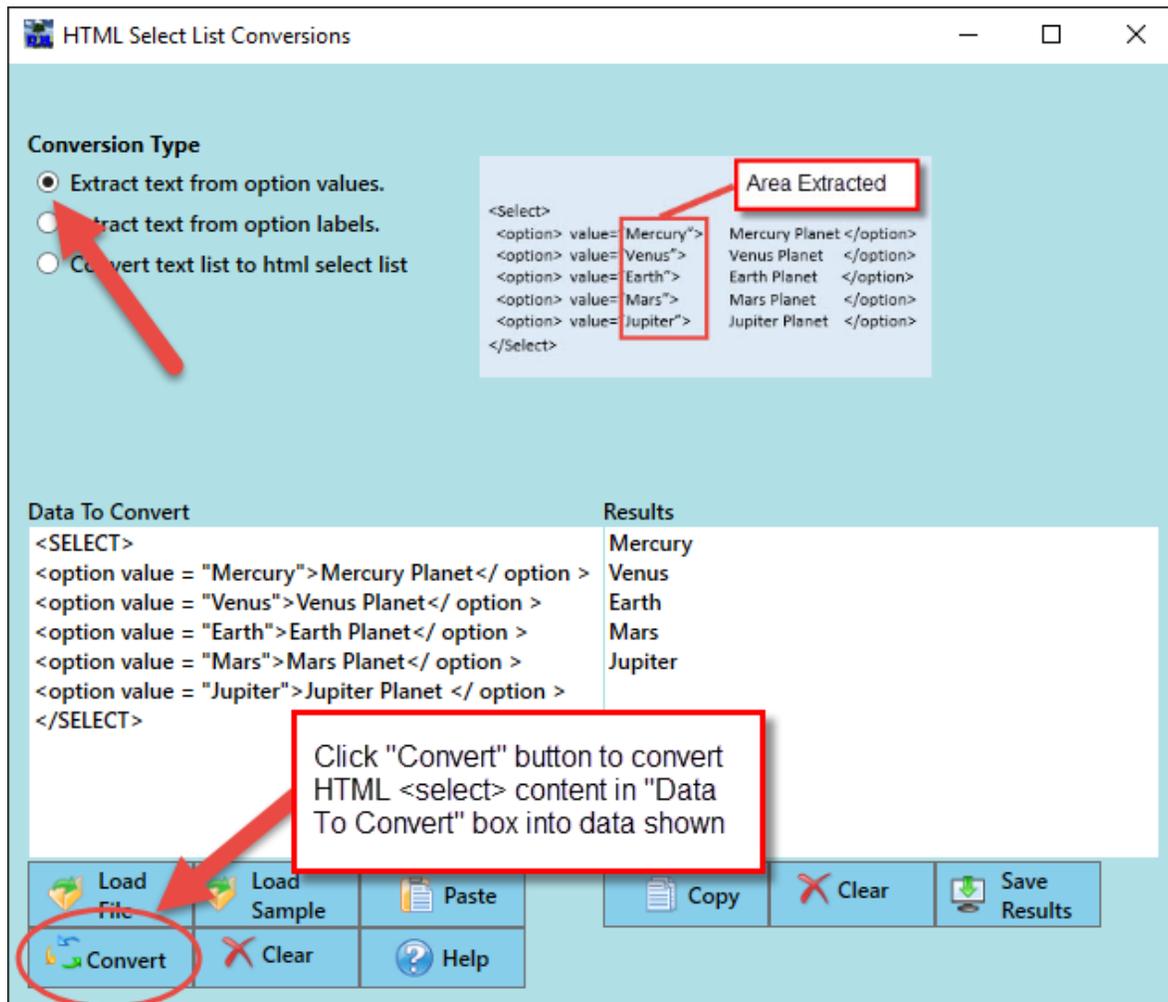


Figure 16 Conversion Type - Extract text from option values

## 1.2 Radio Button – Extract Text from Option Labels

Selecting the radio button Extract text from option labels followed by clicking the “Convert” button results in values being extracted from the block above marked with the letter “B” which would give you the following list in the “Results” box:

- Mercury Planet
- Venus Planet
- Earth Planet
- Mars Planet
- Jupiter Planet

The HTML Select List Conversions window would look similar to the following:

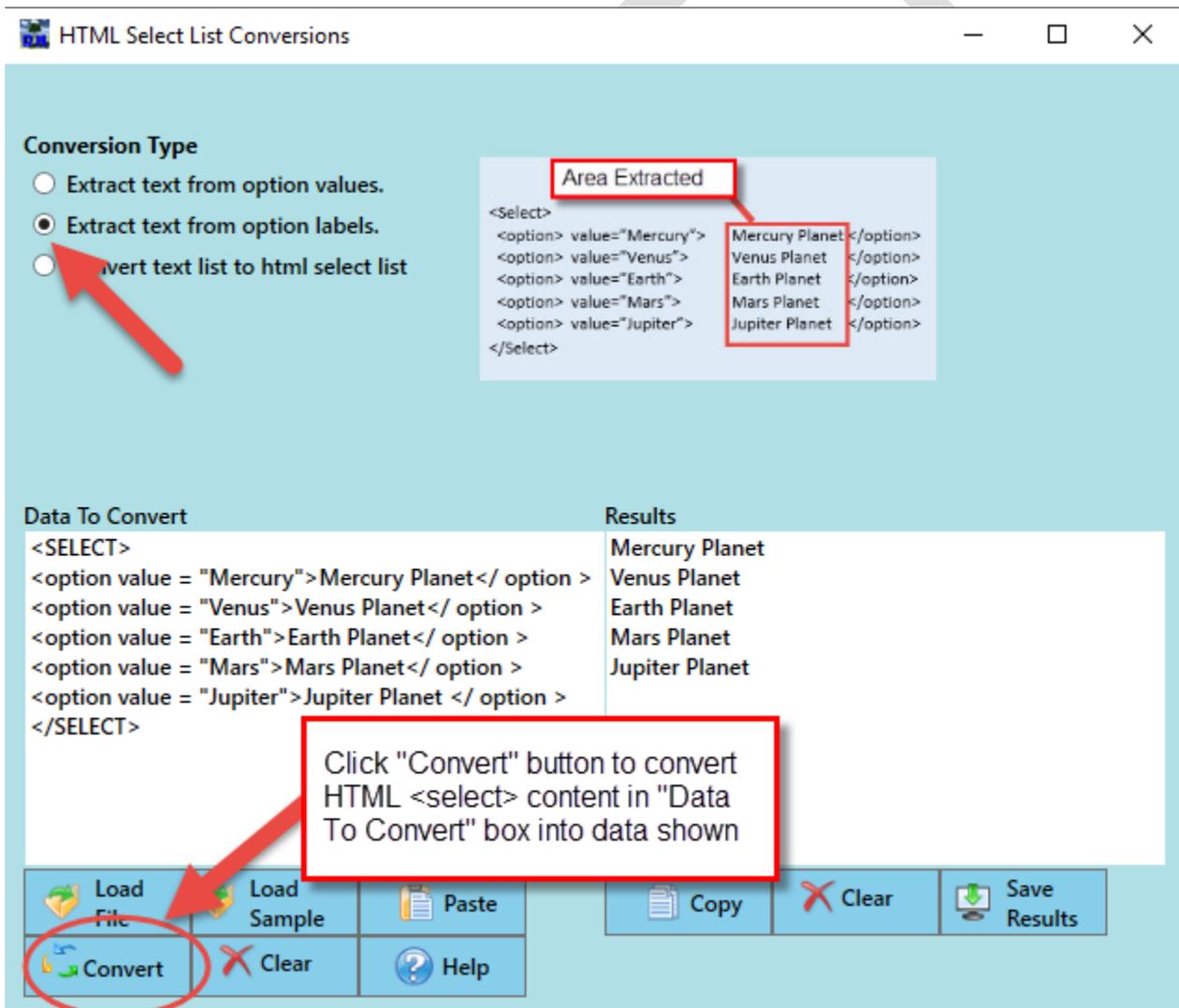


Figure 17 Conversion Type - Extract text from option labels

### 1.3 Radio Button – Convert text list to HTML Select List

Selecting the radio button Convert Text List to HTML Select List will convert a single column of text values in the Data to\_Convert box into a standard HTML <select> list in the Results Window as shown below:

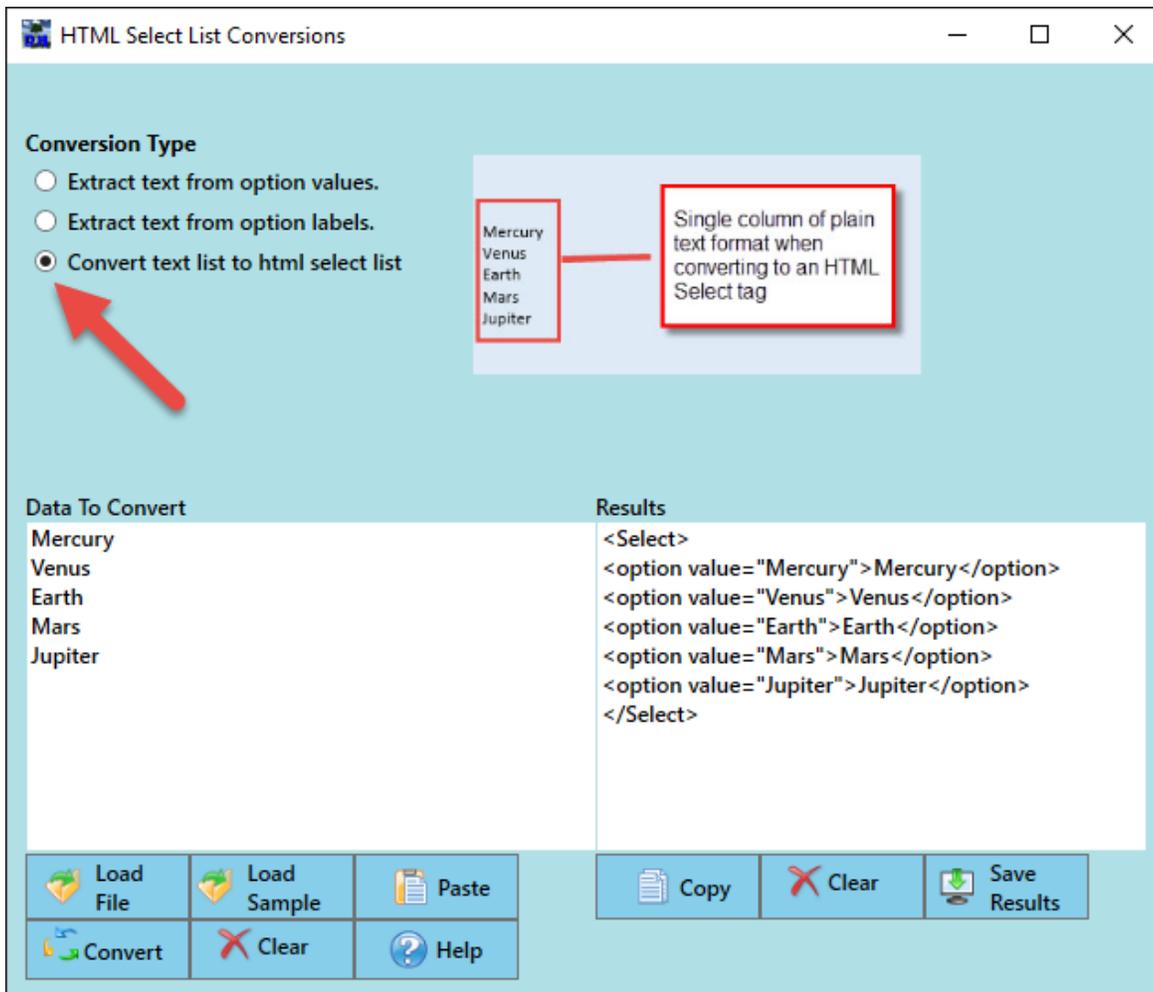


Figure 18 Conversion Type - Convert text list to html select list

The generated HTML <select> statement is very basic. It does not provide a way to have two separate values for the option parameter such as the following:

```
<option value="Mercury">Mercury<Mercury</options>
```

Support for two separate values as shown in the example here along with other options may be available in future upgrade releases to this software.

## 2.0 The Data to Convert Box

The Data to Convert Box is the area you place HTML <select> statements with corresponding <option> statements to be converted. In this initial release of Savvy Realm Web Extractor, the supported <select> statements that can be converted are basic standard formats which have been

shown in the sections above. The following methods may be used to enter standard HTML <select> statements into the Data to Convert Box:

### Available Methods To Enter HTML <select> Statements

- 1) Paste from the clipboard. In this case you may copy the text from a document or displayed <select> list on a web page. If the <select> list you need is not exposed on a web page you may try viewing the HTML page source to locate the <select> and perform a copy and paste from there.

If you are unable to locate the <select> statement by viewing a web page's source, you can try selecting the entire web page source and performing a copy and paste operation. The initial release of Savvy Realm Web Extractor has limited support for extracting <select> list content when reading the entire source of an HTML page.

- 2) Load a text file (.txt) using the Load File Button.
- 3) Perform a drag and drop operation from a text file. If you have a text file with an HTML <select> statement you can drag and drop that file into the "Data to Convert" box from the Windows File Explorer.
- 4) Perform a drag and drop operation from a web browser. Support for drag and drop from a web browser URL address line is limited in this initial release. Part of the problem is due to the same reasons described in option 1 above in the discussion about using copy and paste. At the present time, drag and drop of a web page onto the Savvy Realm Web Extractor is only supported for the following browsers:

- Google Chrome
- Firefox
- Internet Explorer
- Opera

## 2.1 Load File Button

This button allows a user to load a file containing an HTML <select> list to be converted. The below load window will be shown after clicking the button:

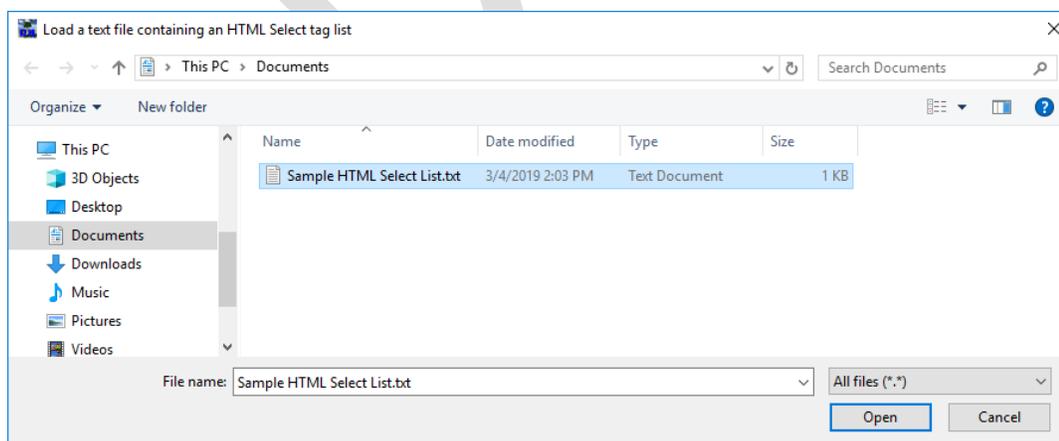


Figure 19 Load Text File

## 2.2 Load Sample Button

This button will load a sample HTML <SELECT> list into the Data to Convert Window. This option is provided so users can see the format supported by Savvy Realm Web Extractor.

## 2.3 Paste Button

The Paste Button will insert previously copied text from the clipboard into the Data to Convert Window

## 2.4 Convert Button

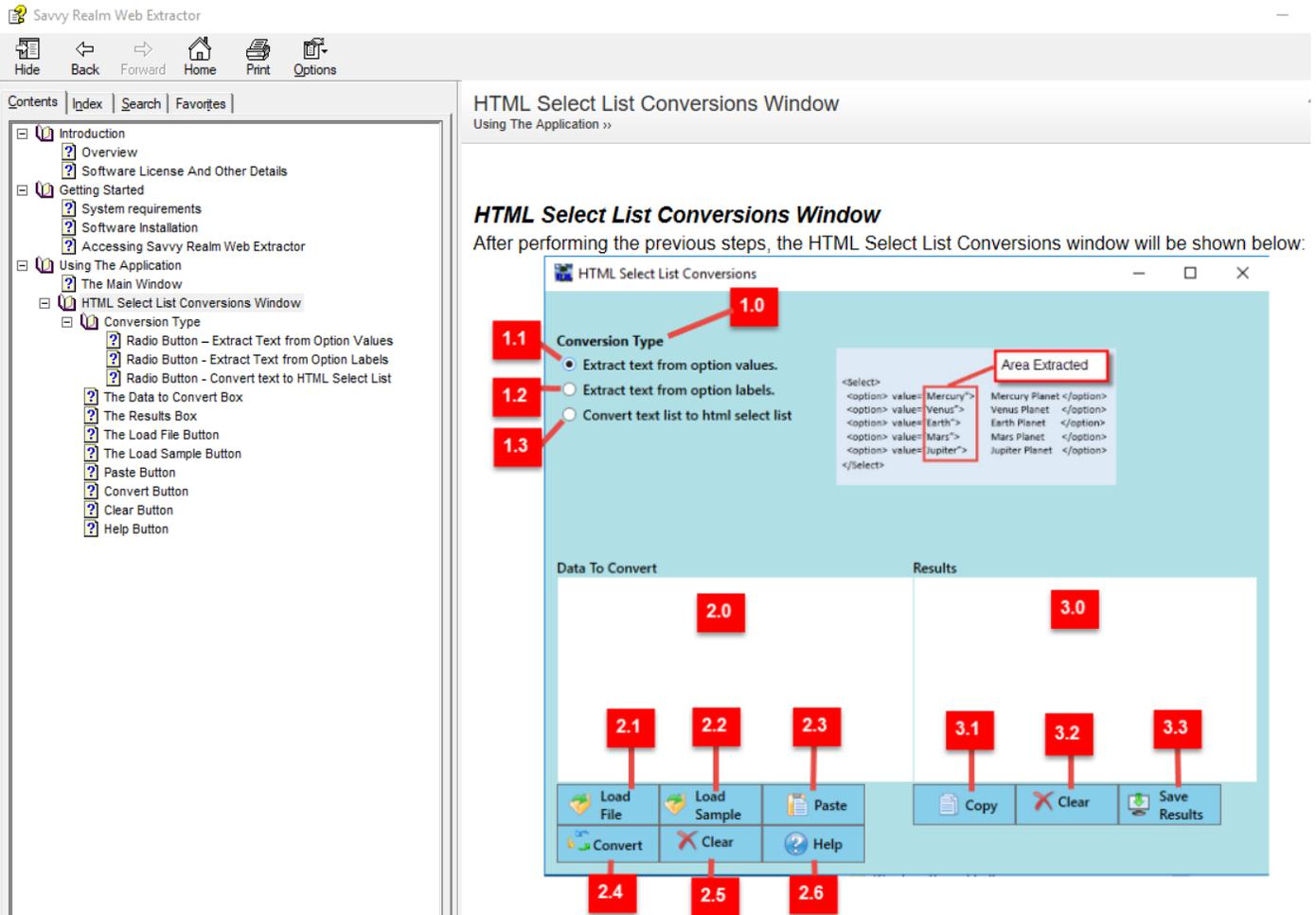
Clicking the Convert Button will start a conversion process to convert content in the Data to Convert Window and store the results into the Results Window.

## 2.5 Clear Button

The Clear Button below the Data to Convert Window will erase all content in the Data to Convert Window.

## 2.6 Help Button

Documentation about how to use the HTML Select List Conversions Window can be obtained by clicking the Help Button on this window. Below is a sample picture of the Windows Help File built for the Savvy Realm Web Extractor:



### 3.0 The Results Box

The Results Box is where content from the Data to Convert box is placed after a conversion has taken place by clicking the “Convert” button.

#### 3.1 Copy Button

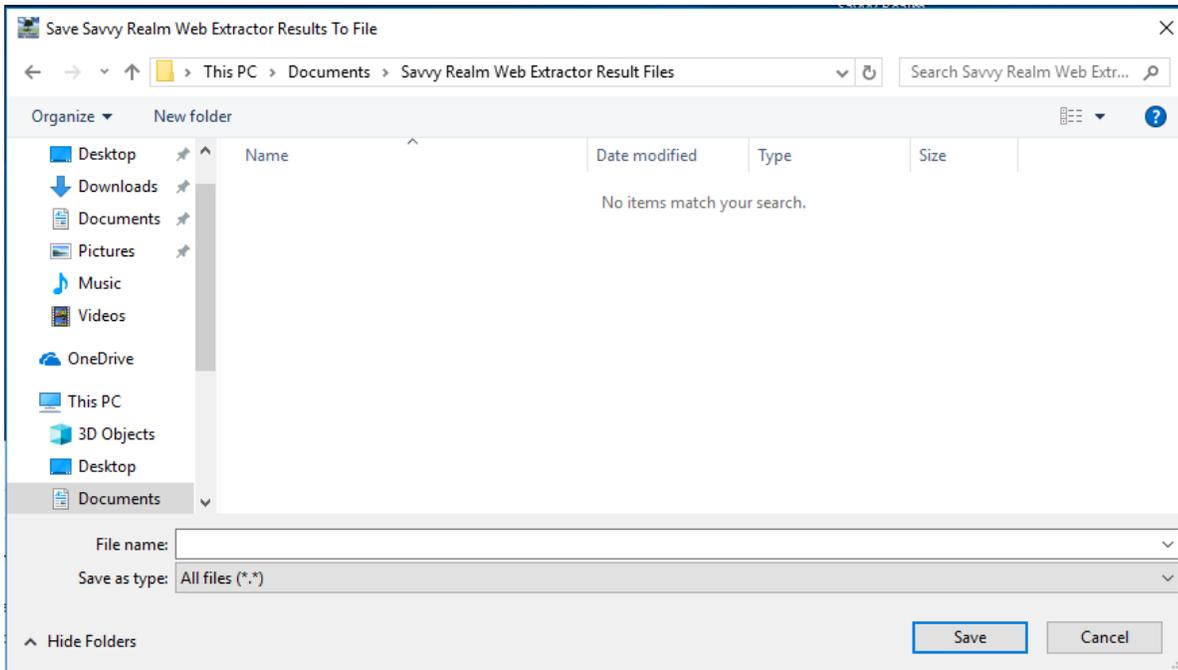
The copy button copies all content in the Results Box and places it into the clipboard so you can paste it into another application.

#### 3.2 Clear Button

The Clear Button located below the Results Box clears all content in the Results Box.

#### 3.3 Save Results

The Save Results button allows the user to save the content in the Results Box. After clicking the button the user will see the following box allowing them to save the Results Box content:



DRAFT