Duplicate Finder Manual

Scan your Lightroom image catalogue to

find suspected duplicates.

Introduction **1**

Requirements 1

Contact and Support 1

Legal Stuff 1

What's Included **2**

New Commands 2

New Smart Collections 3

New Collections 4

Installation **5**

Windows 5

Automatic Installation *5*

Manual Installation *5*

Mac 5

Setting up Lightroom 6

Installing ExifTool *6*

Enter Your Licence Key *7*

Searching **9**

Excluding Photos 10

Running the Search 11

Finding your Results 12

Limiting Results 13

Deleting Unwanted Photos **15**

Introduction

This is a straightforward plug-in which will attempt to find duplicate images in your Lightroom catalogue using the EXIF data associated with each photo.

All photos that are potential duplicates are shown. The decision on which ones to keep or discard is yours – this is too dangerous to leave to an automated process! There are easy methods of reducing the amount of work that needs to be done after isolating duplicates using the plugin.

You can improve the accuracy of the scan by installing the free "EXIFTool" by Phil Harvey, from https://www.sno.phy.queensu.ca/~phil/exiftool/install.html .

Requirements

- Lightroom 4 or newer, including CC versions.
- Mac or Windows.

Contact and Support

You can get support or ask questions in the support desk. Please check the FAQ before contacting support, many common questions are already answered here.

https://www.lightroom-plugins.com/support/index.php?/Lightroom/Troubleshooter/List

https://www.lightroom-plugins.com/faq.php

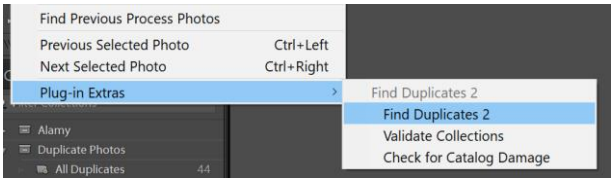
Legal Stuff

This plug-in is not specifically endorsed by Adobe Inc®.

What's Included

New Commands

Three new commands have been added to the "`Library`" menu in Lightroom:



Please note that there are several different "Plugin Extras" menus in Lightroom. The new commands for this plugin are added to the "Library" menu and nowhere else.

FIND DUPLICATES

This performs the scan of the catalog, marking any suspected duplicates as it goes. Each potential duplicate is added to a number of new collections which categorise them by "largest", "newest" and other groups. They are also added to a static collection grouped together so that you can see each set in turn, if you need to review the final results.

How it works is fully described below under "Searching".

New Smart Collections

To make it easier to find your images, there are several new Smart Collections provided. Smart Collections automatically find images in your catalog without having to be manually refreshed when something changes. These new Smart Collections are grouped together under a single Set called "Duplicate Photos" and are all created automatically the first time you run the plugin.

ALL DUPLICATES

Shows all images which have been identified as being potentially duplicates. The plugin doesn't attempt to work out which one you might want to keep so all matches are shown in this collection. Please remember that the plugin has no idea of which one of many similar photos you may want to keep.

FILE – LARGEST

This contains, for each set of duplicate photos, the one with the largest file size. If any of the duplicates are currently offline or missing, they will not be included here since it is not possible to get the file size of a missing file.

FILE – SMALLEST

This contains, for each set of duplicate photos, the one with the smallest file size. If any of the duplicates are currently offline or missing, they will not be included here since it is not possible to get the file size of a missing file.

MEGAPIXELS – LEAST

This contains, for each set of duplicate photos, the one with the smallest number of megapixels.

MEGAPIXELS – MOST

This contains, for each set of duplicate photos, the one with the highest number of megapixels.

DATE – OLDEST

This contains, for each set of duplicate photos, the one with the oldest "Date Time" value. This is different to the "Date Time Original" value and is typically the date on which this copy of the photo was created.

DATE – NEWEST

This contains, for each set of duplicate photos, the one with the newest "Date Time" value. This is different to the "Date Time Original" value and is typically the date on which this copy of the photo was created.

EDITS – LEAST

This contains, for each set of duplicate photos, the one with the least number of edits recorded by Lightroom. This value, while useful, may not be completely reliable because of the way that Lightroom records edits. A long sequence of edits may be recorded as a single change if they happen quickly.

EDITS – MOST

This contains, for each set of duplicate photos, the one with the highest number of edits recorded by Lightroom.

New Collections

A standard collection is not updated automatically by Lightroom – it keeps the same set of photos until they are manually changed. The collections listed here are cleared and recreated by the plugin each time you run a new search.

ALL DUPLICATES (MATCH ORDER)

Each of the files identified as a potential duplicate is added to this collection. By setting Lightroom's sort order to "Custom Order", you can see each of the duplicate sets in the order that they were discovered. This may help in the case where you have not used the "Date/Time Original" value in the duplicate search, and no other item suitable for sorting exists.

Installation

This plug-in adds a new field to your Lightroom database. This is used to store a number indicating which images are seen as duplicates. Each set of duplicates will have a different number.

Before you install this, you should make sure you have a backup of your Lightroom catalog. If you do not already have automatic backups enabled, you can perform a single backup by going to "`Edit→Catalog Settings...`" and changing the "`Back up catalog`" list to "`Next time Lightroom starts only`", and restarting Lightroom.

Windows

Automatic Installation

The easiest way to install on Windows is to use the MSI installer. Simply double-click and run through the prompts that appear.

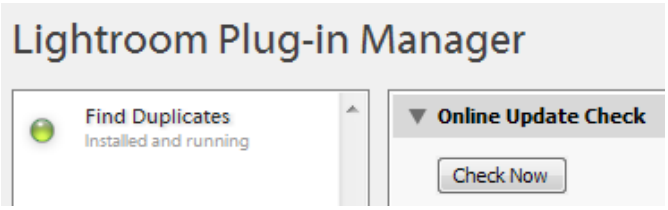
Manual Installation

If you prefer to install manually, download the ZIP file instead of the MSI. When it has downloaded, follow the instructions below for installing on the Mac; the steps are identical on Windows.

Mac

1. Download the ZIP and extract it to your desktop. You should see a folder called "**`FindDupes2`**" (Mac) or "**`FindDupes2.lrplugin`**" (Windows), containing the plug-in.
2. Move this folder anywhere you like.
3. Open Lightroom's Plugin Manager from the File menu, and click the "Add" button at the bottom of the list.
4. Use the file dialog that appears to select the plugin (ending in **`.lrplugin`**).

It should now be listed in the Plugin Manager as "`Installed and Running`".



Setting up Lightroom

There is nothing more you need to do to prepare Lightroom for using the plugin. The Smart Collections will be created the first time you run a duplicate search.

Installing ExifTool

This is an optional step, only required if you wish to match timestamps with sub-second accuracy or match other fields which Lightroom does not support itself.

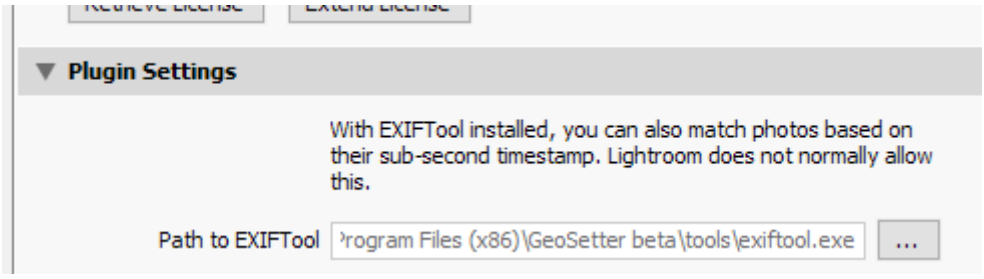
Since July 2016 it has been possible to search on additional fields that are not provided by Lightroom, by installing Phil Harvey's free "exiftool" program. This is most often used to expand the date/time search to be aware of sub-second times; Lightroom only tracks times to the nearest second, and this can result in duplicates being identified in sequences of photos shot at high speed. It is also capable of extracting large amounts of information from most image types and can be very useful in identifying high-speed bursts or regular/HDR pairs where Lightroom simply doesn't store any information that shows this.

To install ExifTool, download the appropriate package from http://www.sno.phy.queensu.ca/~phil/exiftool/ (either the "Windows Executable", or the "Mac OS X Package").

On Windows, extract the zip archive and place the files wherever you want them on your PC. You may need to rename "exiftool(-k).exe" to "exiftool.exe"; if this is the case then the plugin will warn you.

On the Mac, install the program using the installer contained in the dmg file you downloaded.

Next, open Lightroom's "Plugin Manager" and select the "Find Duplicates 2" plugin. In the Plugin Settings panel, click the button to select the location of exiftool if it has not been found automatically.

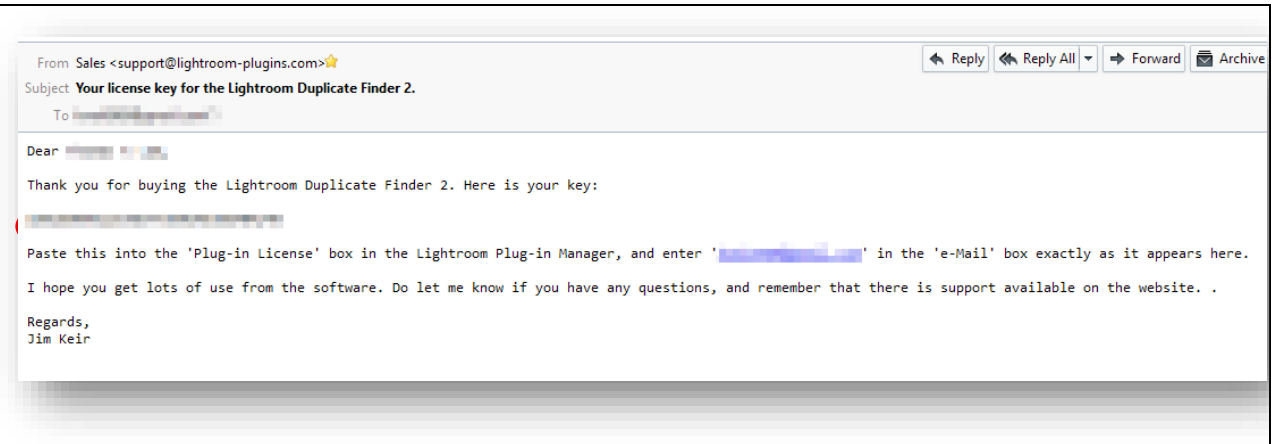


With exiftool correctly installed, more options will be available during the search. The penalty is that exiftool must scan your photos one by one during the search, which can take some time.

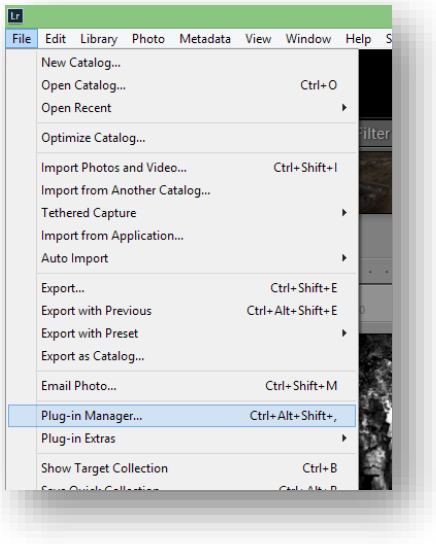
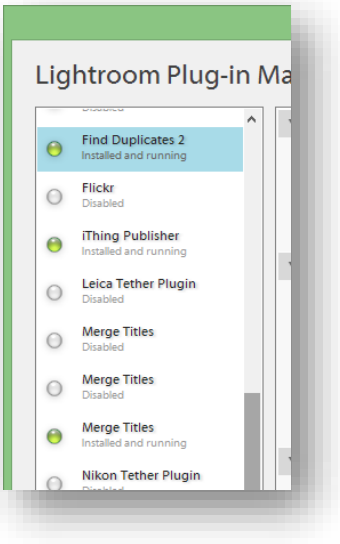
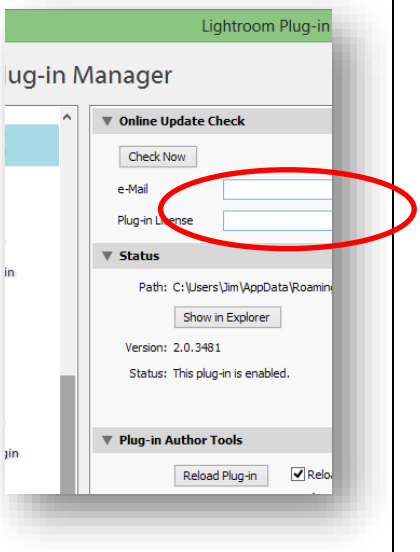
Enter Your Licence Key

If you have bought the plug-in, you should receive a licence key in an email. Without this, the plug-in will stop after finding a small number of duplicates. To enter the licence key, simply copy and paste both the key and your e-mail address from the email to the box in the Plugin Manager when the Find Duplicates plugin is selected.

The Plugin Manager can be found on the File menu. If you would like a more detailed tutorial on how to use this, or on other aspects of Lightroom usage, please visit https://www.learn-to-lightroom.com .



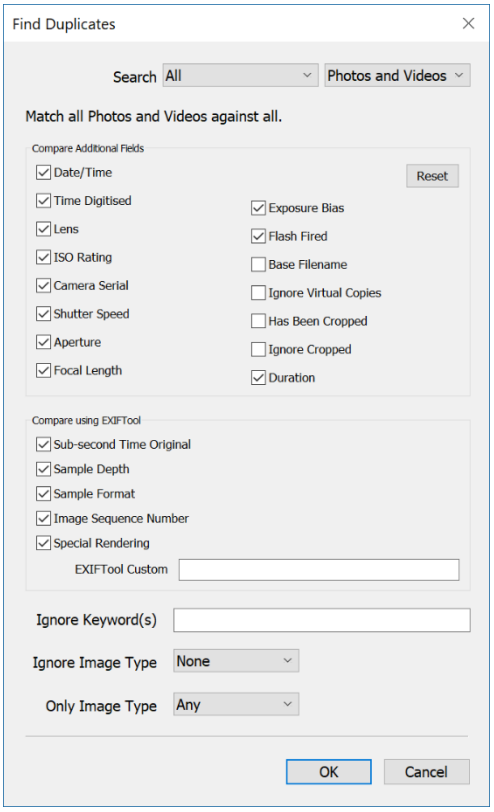
1) Find your license.

		
Open the Plugin Manager.	**Select the correct plugin.**	**Paste in your email and license key.**

SEARCHING

This assumes you already have some images stored in Lightroom.

Select the "`Library→Plug-in Extras→Find Duplicates 2`" menu item. A dialog will appear:



The plugin can search in one of three ways:

- **ALL PHOTOS**

Your entire catalog is searched for duplicates.

- **WITHIN SELECTED**

Only duplicates within the currently selected photos are found.

- **FOR SELECTED**

Duplicates of the selected photos are searched for in the entire catalog.

- **WITHIN DISPLAYED**

Essentially the same as "Within Selected", but automatically selects the entire current set of displayed photos.

Next you can choose what EXIF data to compare, to determine if one photo is a duplicate of another. The default values will work in almost all cases, and can be restored using the "Reset" button.

A limitation of basic EXIF data is that the timestamp is limited to one-second accuracy. This means that photos taken very close together with no settings changes can be picked up as duplicates. If you are bracketing then usually the exposure or aperture will change, so this can be used to filter these out. Photos taken this close together will usually appear to be very similar and are still probably worth seeing as duplicates, but if you want to keep a high-speed sequence like this then you can mark them as "not duplicates" permanently using the "Not a Duplicate" tickbox in the Library module's metadata panel.

Alternatively you can install Phil Harvey's ExifTool as described in Installing ExifTool. This bypasses the limitations of Lightroom and allows times to be compared with sub-second accuracy.

The other options all affect which photos are included in the results, allowing you to tailor the search to your own requirements. When ticked, the options like "Aperture" and "Exposure Bias" require those values to be matched before two photos are considered duplicates.

Excluding Photos

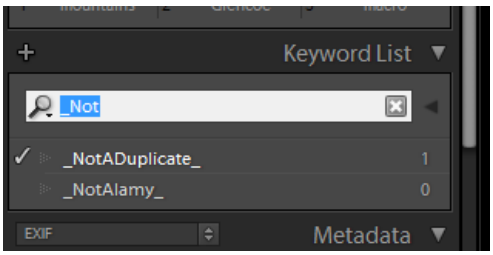
Virtual copies are handled slightly differently. With "Ignore Virtual Copies" ticked, no virtual copies will be included in the results.

The "Ignore Image Type" and "Only Image Type" options restrict the search to exclude or include the specified type.

If you want to permanently exclude a specific photo from the search, you can do this in two ways. Firstly, you can set a keyword on the photo. This does not necessarily mean that the keyword will be visible if you export the photo.

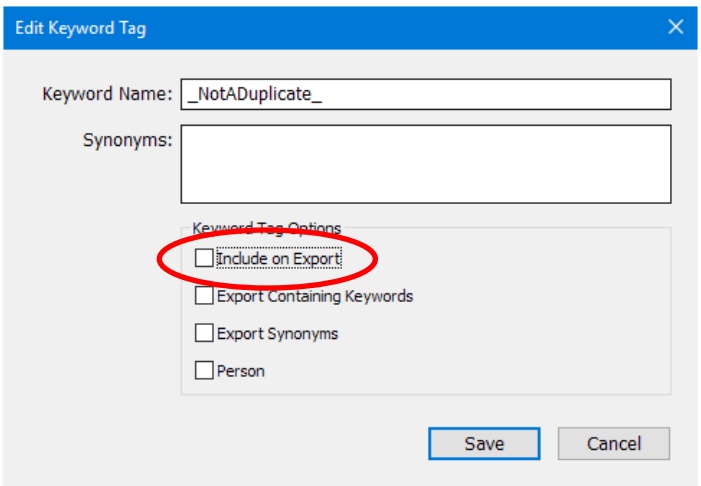
In the Library module, pick the photo or photos you want to be excluded. Add a keyword of your choice, for example "_NotADuplicate_".

If this is the first time you have used this keyword, you can mark it as non-exportable. This means that if you export a copy of the photo, maybe uploading it to a photo agency or web service, the keyword will not be included in the exported photo. Still in Lightroom's Library module, expand the "Keyword List" section and search for the keyword you just created.



Select the "_NotADuplicate_" keyword from the list, and then right-click and choose the "Edit Keyword Tag" option from the menu. In the dialog that appears, clear the "Include on Export" checkbox.

The plugin does not care whether it is included in export or not.



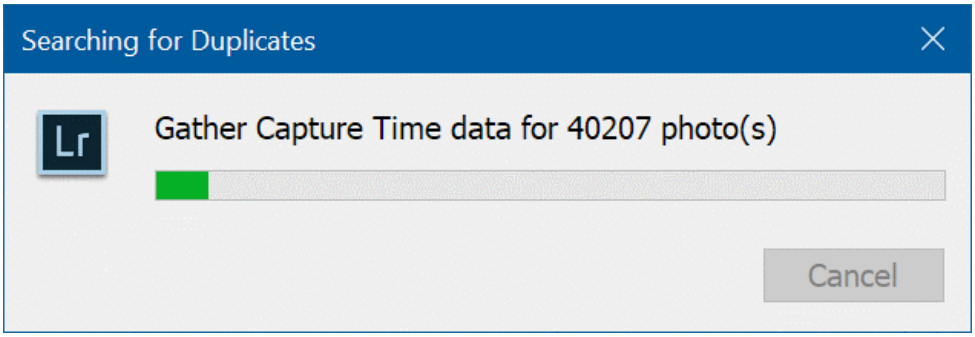
Running the Search

The Duplicate Finder is available in Lightroom's Library module, on the Library menu. To run a search, first decide which photos you wish to be included.

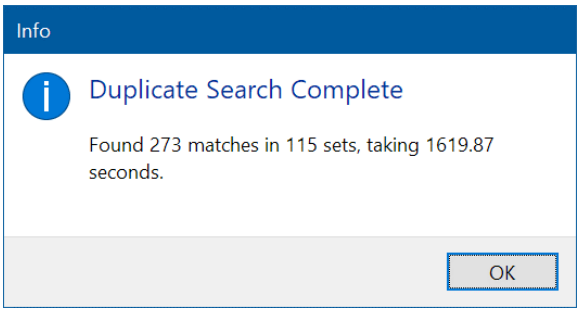
If you are going to search the entire catalog, then you don't need to select any photos beforehand. For both of the "For Selected" and "Within Selected" search modes, you need to make sure that you have selected the correct photos before you run the search. "Within Displayed" automatically makes the selection for you.

To begin the search, go to Lightroom's Library module, then select the "Library->Plug-in Extras->Find Duplicates 2" menu item.

A progress window will appear:



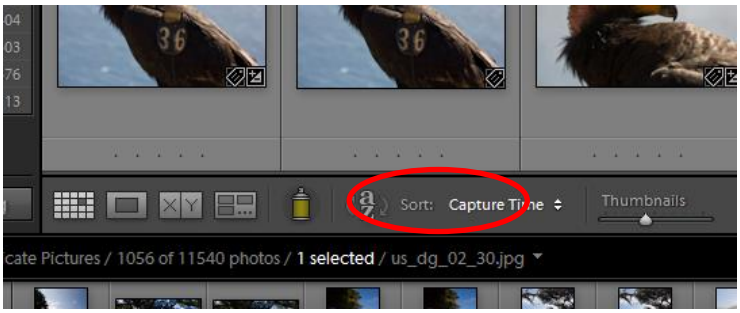
Once the search has completed, a summary is shown:



FINDING YOUR RESULTS

Once a search has completed, you can view the results. The search has filled in a new field in the database called "Duplicate Id", and it is this that the Smart Collection searches for. Any photo which is considered to be a duplicate will have a number in this field, the same number being assigned to each duplicate of that photo.

To view the results, simply select one of the collections under the "`Duplicate Photos`" Collection Set. Normally the easiest way of seeing the duplicate sets is to sort the main display by "Capture Time". However, if you didn't use the capture time as part of the search, then this won't work. Instead, select the "All Duplicates (Match Order)" collection. This collection is built by the plugin each time you run it, and contains all of the matched photos in the order in which they were matched. You can now order the grid display by "Custom Order", and the duplicates will be shown grouped together.



It's now up to you to decide what to do with your duplicates. You can, of course, use Lightroom's normal Library Filter tools to limit the selection, for example by file extension, so that you can delete unneeded photos.

LIMITING RESULTS

With the "Found Duplicates" Smart Collection displayed, you can apply any combination of normal Lightroom filters to help you isolate the copies you want to keep from those you want to get rid of. The plugin deliberately makes no attempt to make this decision for you! From the online FAQ:

> *The plugin will identify all images that are likely to be copies of each other, based on the image metadata. Many people ask why it doesn't automatically delete duplicates, or why it doesn't exclude the original.*
>
> *This is simply down to safety. For a commercial photographer, losing a valued original image may be disastrous. Would you really want a bit of software deciding which photo to delete based on some kind of pre-written rule?*
>
> *Suggestions have included:*
>
> - *Keep the one with the shortest name, assuming that duplicates will have a number or "copy of" in the name. So, we delete "Photo (Original).jpg" and keep "Photo (Web).jpg". Wrong.*
> - *Delete JPEGs, keep the RAW. But what if you shoot in JPEG? What if you convert to DNG and use that as the master, and the RAW is the one you would want to get rid of?*
> - *Keep the oldest one, on the assumption that the oldest is the master. So, we delete the one to which you have made loads of production edits and keep an early preview you exported for your client. Wrong.*
> - *Keep the newest one, on the assumption that it will be the one with the most work done to it. So, we delete the older original and keep the web-sized, watermarked export you prepared for a customer. Wrong.*
> - *Keep the largest one, on the assumption that all smaller ones are previews. So, we delete an original scanned copy and keep a resized-for-printing copy with a banner across it. Wrong.*
>
> *In short, no matter what rules you have for identifying which image is the "original" and which are "duplicates", there will be someone else who does the opposite. The plugin will help you to **identify** which images to look at by eliminating most of your library, and letting you display only the potential duplicates next to each other. You can then use Lightroom's own tools to identify, mark and delete any that you don't want, according to your own rules.*

In short, every person will define “original” as something different. Rather than take the risk of automatically identifying the wrong copy of a photo for deletion, the plugin leaves this up to you.

For a filtering tutorial, visit https://www.learn-to-lightroom.com .

Another way of limiting your results is to use the other Smart Collections, named “Least/Most/Largest/Smallest” etc. Each of these collections is guaranteed to only include at most one photo from each set, so if you’re absolutely sure that you always want to delete photos with the least megapixels, you can display that collection, select all and reject.

In many cases, if you have duplicates they will be in different folders, or have different names, or have some characteristic in common that lets you identify many photos at once. You should rarely if ever have to trawl through all duplicate photos, rejecting them one by one.

Deleting Unwanted Photos

Once you have limited the photos in your view, it is possible to delete unwanted photos. If you try and delete from the Smart Collection, Lightroom will warn you that this is not possible. It assumes that you are trying to remove a photo from the collection, and since this is a Smart Collection which is populated automatically, this is not possible.

Removing photos is still easy though. You need to use Lightroom's "Reject" feature. With this, you can reject photos whether they are part of a collection or not. The easiest way is to press X on the keyboard, but as usual with Lightroom there are many ways of using the same feature. Once you have rejected some photos, you can finally delete them all at once using the "Delete Rejected Photos" command on the "Photo" menu.

For this to work, you must be viewing an item in the "Folders" panel, or the "All Photographs" item in the "Catalog" panel. If you're viewing anything in the "Collections" panel, including the "Duplicate Photos" collections, then the "Reject" feature won't work. Lightroom sees this as an attempt to remove rejected photos from the *collection*, not the *catalog*. Removing a photo from a Smart Collection is impossible, because the contents are determined automatically from the collection's definition. Removing a photo from a standard Collection *is* possible, but the photo will remain in the catalog.