



## Installing and Running a Data Quality Loqate Standard Address Application

This use case describes how to install and run a Data Quality (DQ) Loqate Standard Address application. It includes the following topics:

- [Prerequisites](#)
- [Installation](#)
- [Downloading the Address Data](#)
- [Configuring the Data Quality Runtime](#)
- [Configuring the Data Quality Scoring Plan](#)
- [Verifying Loqate Data](#)
- [Checking the Expiration Status for Loqate Data](#)
- [Additional Resources](#)

---

### **Prerequisites**

This section summarizes the prerequisites that are required before using the DQ Loqate Standard Address application. A basic knowledge of building and running DQ plans is assumed.

The DQ Loqate Standard Address application requires:

- iWay Data Quality Server (DQS) GUI or iWay Integration Tools (iIT) with DQ plugins.
- A Loqate license.
- A sample DQ plan, which you can download from:

[http://techsupport.ibi.com/iway\\_patch/Loqate/LOQATE\\_SCORING\\_2016.zip](http://techsupport.ibi.com/iway_patch/Loqate/LOQATE_SCORING_2016.zip)

- The latest Loqate installation, which can be downloaded from:

<ftp.loqate.com>

To obtain the login credentials (user name and password) for this FTP site, contact Customer Support.

Once you have logged in to this FTP site, you must navigate to the following folder to download the executable files:

*/2017Q1.0/AppInstaller*

**Notes:**

- The standard installer for Windows is *setup\_64.exe*.
- If you want to download the Coding Accuracy Support System (CASS) version, then download *setup.exe* (32-bit version).
- If you want to download the Linux version, then download *loqate\_installer\_x86\_64*.

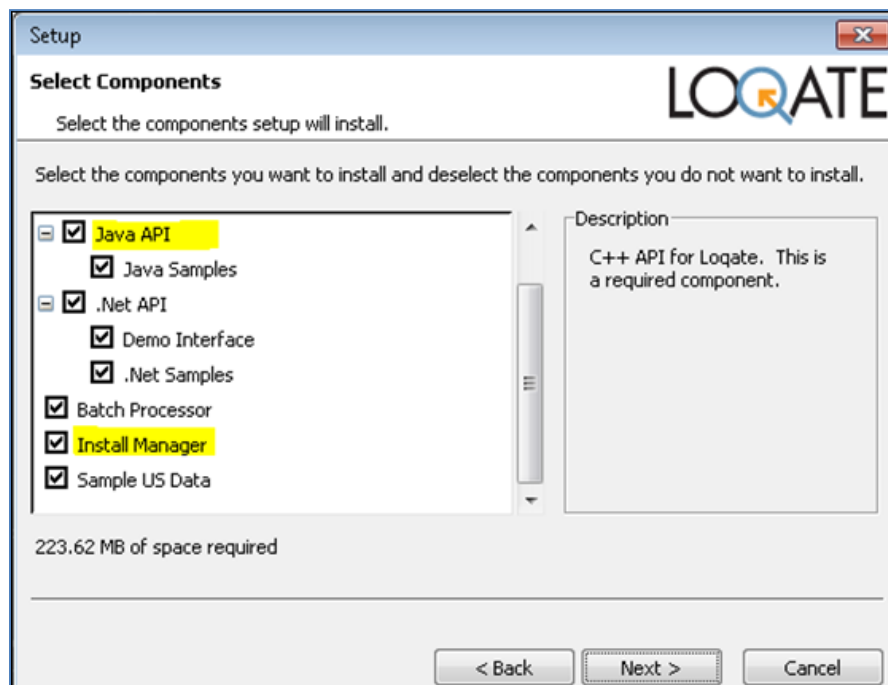
**Installation**

This section describes how to install DQ Loqate.

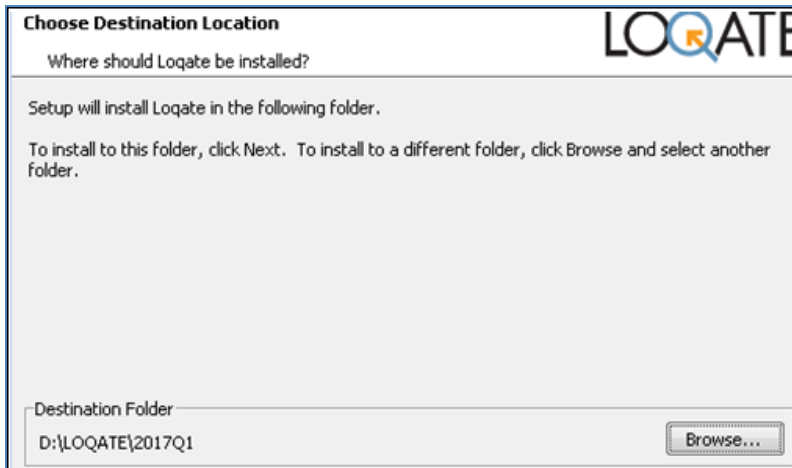
1. Run *setup\_64.exe*.

This installs most of the components, but ensure the following check boxes for these components are selected:

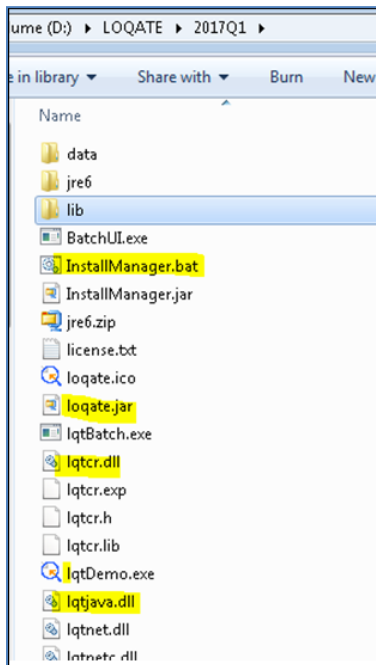
- *Java API* (Loqate .jar and .dll's) which will be the 64-bit version for the standard (non-CASS) version. For CASS, you *must* instead run *setup.exe*, which is the 32-bit version.
- *Install Manager*, which you will use to download your address databases, also known as Global Knowledge Repositories (GKR).



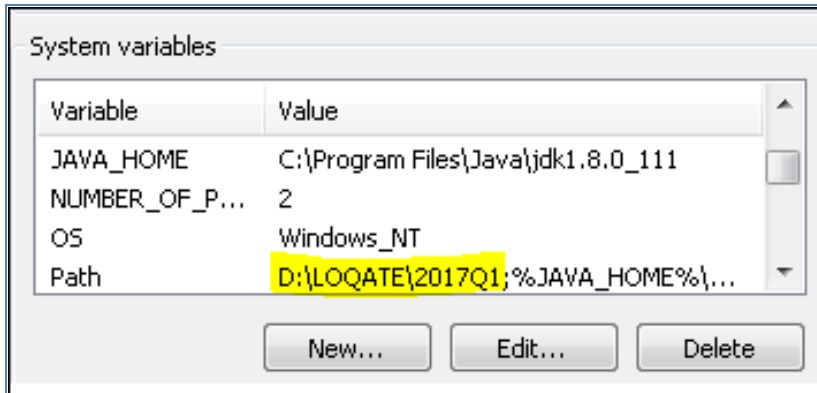
2. Click Next.
3. Select the directory where you wish to install DQ Loqate, for example, D:\LOQATE\2017Q1, as shown in the following image.



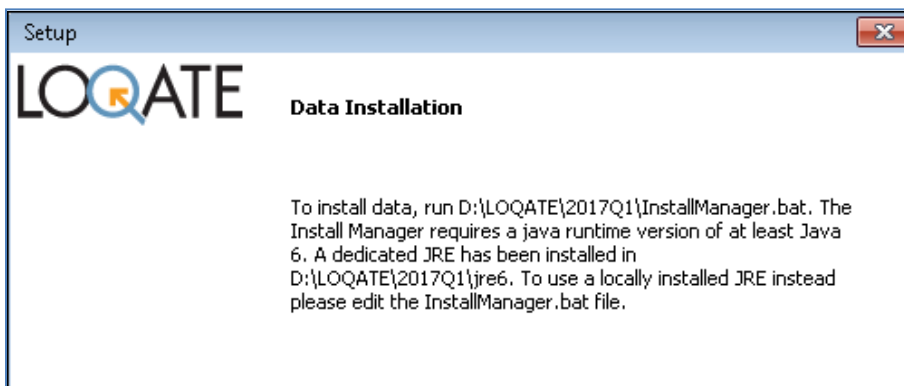
Once you have completed this step, the following key files should appear in the list, as shown in the following image.



Note that now your Windows System Variables Path has automatically been updated to include your Loqate directory, as shown in the following image.



This is required so that any program (for example, DQ) can find the Loqate .dll files.

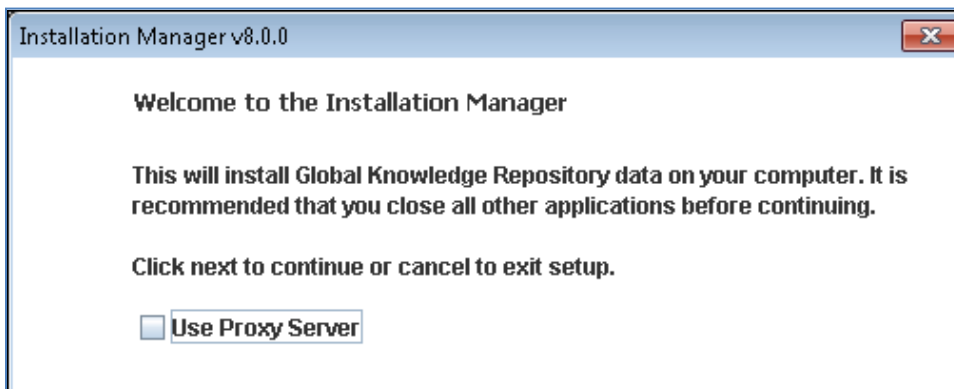


### **Downloading the Address Data**

This section describes how to download the Address Data, as mentioned in the installation process.

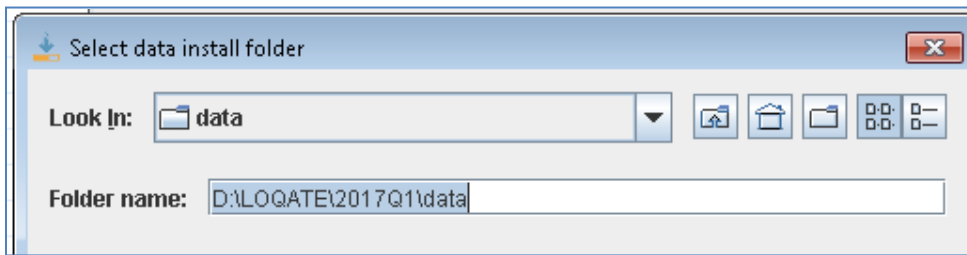
1. Run *InstallManager.bat* to download your databases.

The following Welcome message appears.



2. Click *Next*.

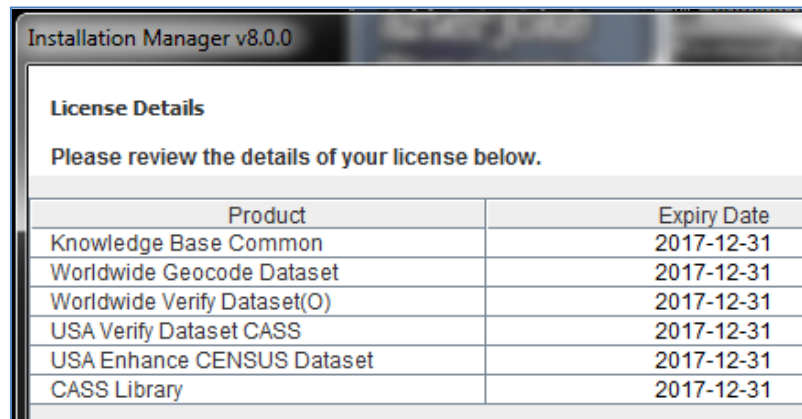
3. Select the location where you installed DQ Loqate (for example, D:\LOQATE\2017Q1), and create a data folder. This will be where the data files will be installed, as shown in the following image.



4. Enter your Loqate License key.
  - Customers are provided a Loqate License key by iWay licensing/support, after a Loqate License Request form is submitted and processed.
  - Information Builders employees can go to the following internal network site for a Loqate License key:

<http://eimwiki.ibi.com/mediawiki/index.php/Category:Loqate>

After the license server is contacted (which requires internet access), a list of your entitlements will be displayed, as shown in the following image.

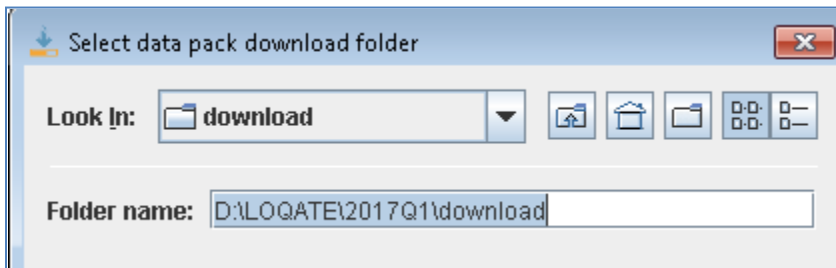


License Details	
Please review the details of your license below.	
Product	Expiry Date
Knowledge Base Common	2017-12-31
Worldwide Geocode Dataset	2017-12-31
Worldwide Verify Dataset(O)	2017-12-31
USA Verify Dataset CASS	2017-12-31
USA Enhance CENSUS Dataset	2017-12-31
CASS Library	2017-12-31

If you get an error stating that the license is invalid, it may be due to a network, environmental, or security issue. For Windows10, you can try running the InstallManager.bat file as an Administrator.

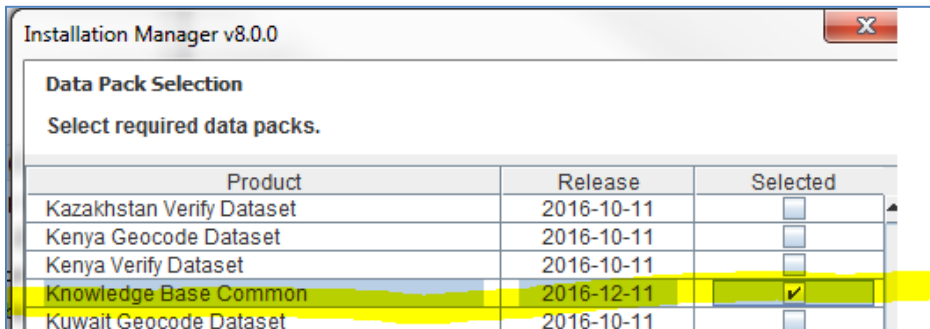
5. Select the directory where you wish to download the datapacks, which are zipped files of the address database.

You can download the datapacks to any directory you wish (for example D:\LOQATE\download), as shown in the following image.

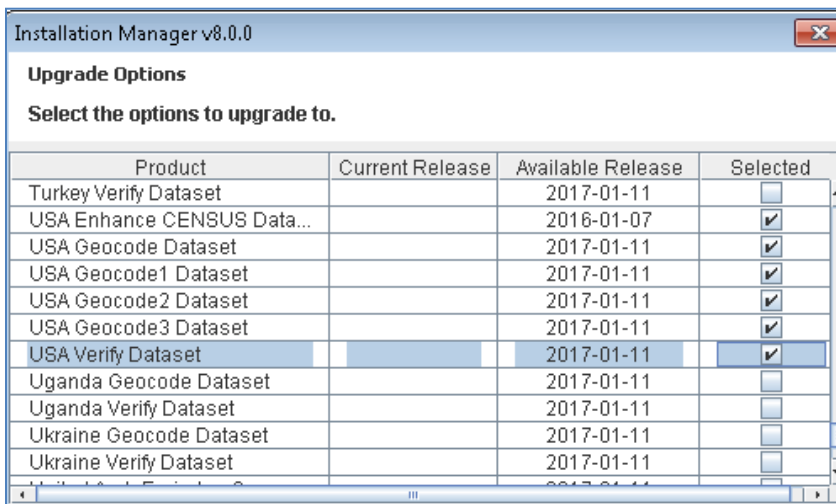


The datapacks will then be uncompressed and installed into the \data folder.

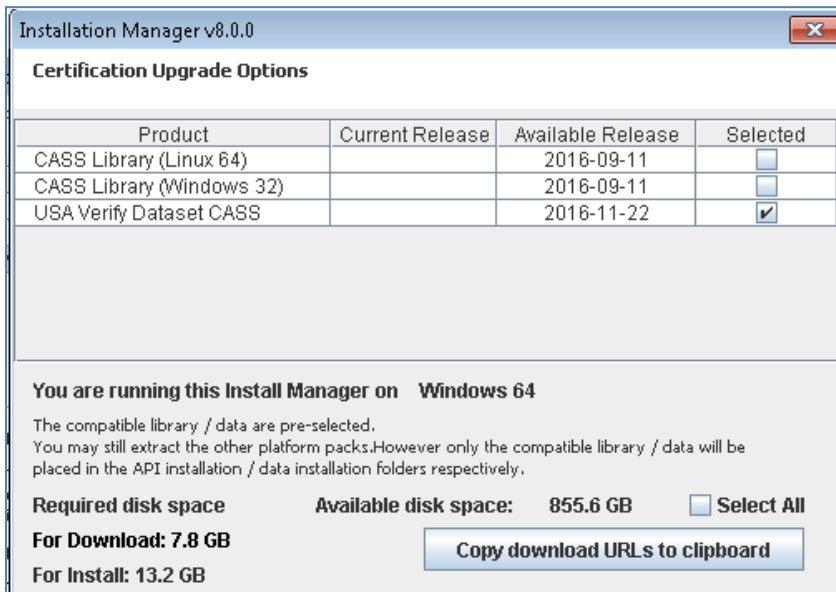
6. Select *Knowledge Base Common*, as shown in the following image.



A list of all your licensed countries appears, as shown in the following image.



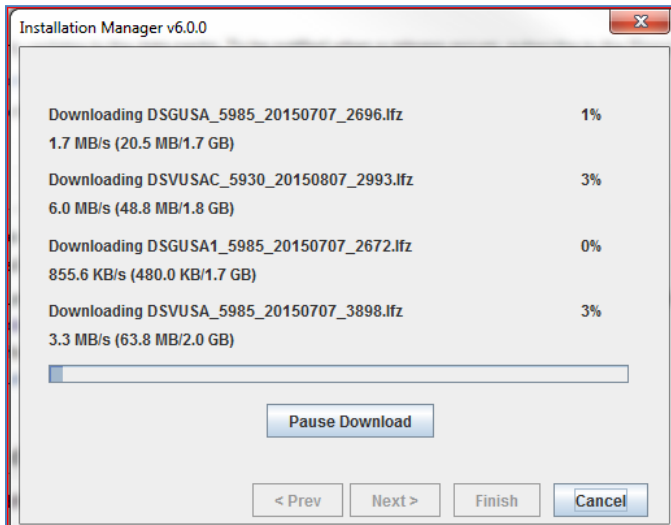
7. Click the *Select All* check box or select the individual products that you want to download at this time.



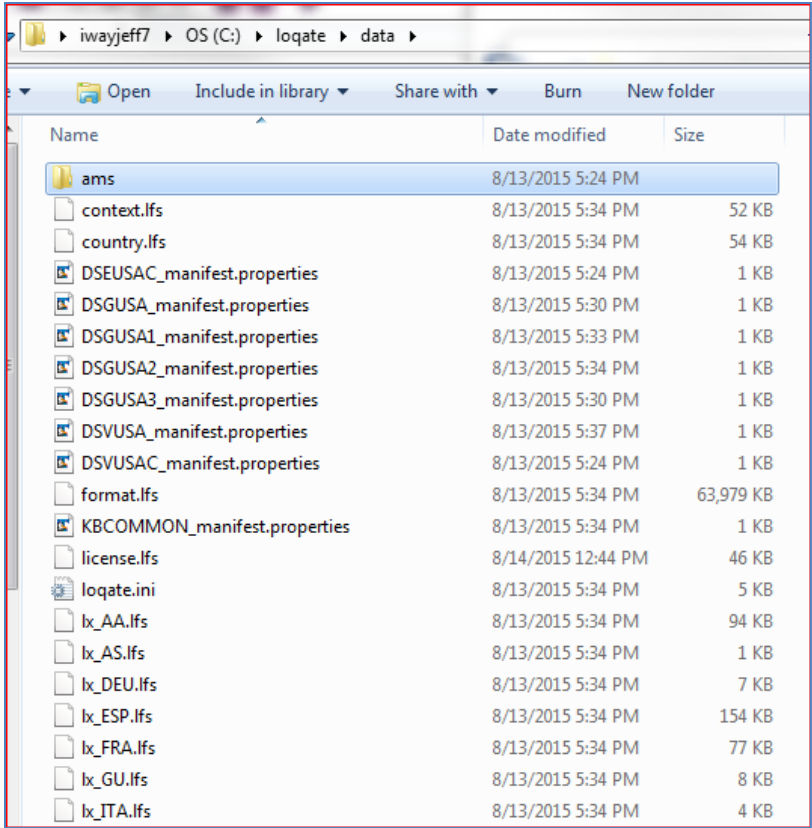
**Note:** If you already have datapacks in your download directory, you may not be presented with an option to download them again (if they are already up to date). If you want to override this behavior, delete the existing datapacks (.lfz files).

8. Click *Next*.

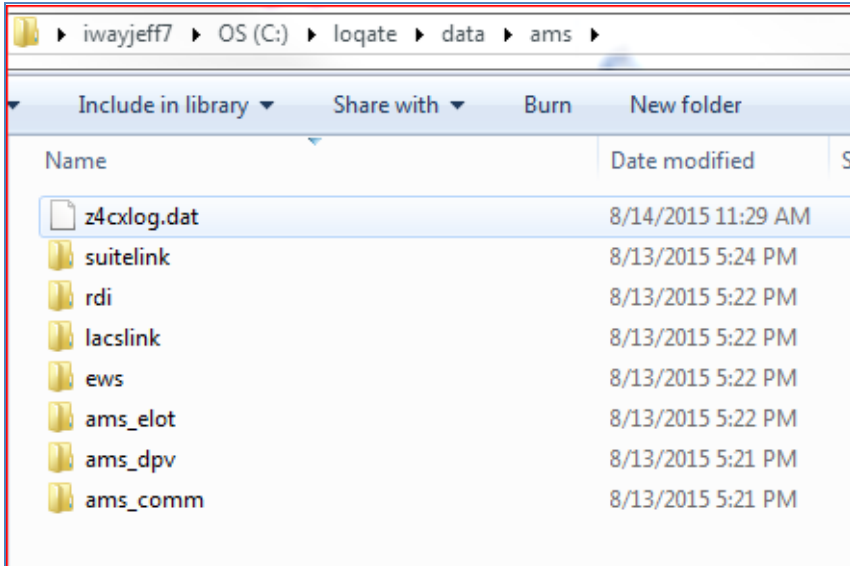
The downloads begin, as shown in the following image.



When the downloads are finished, your data directory should contain numerous .lfs files, as shown in the following image.



If you have licensed and installed US/CASS, it will be in an AMS subdirectory (not required for this use case).



If you wish, you can remove the downloaded datapack.lfs files to save space (a complete US with geocode and CASS is about 8GB).

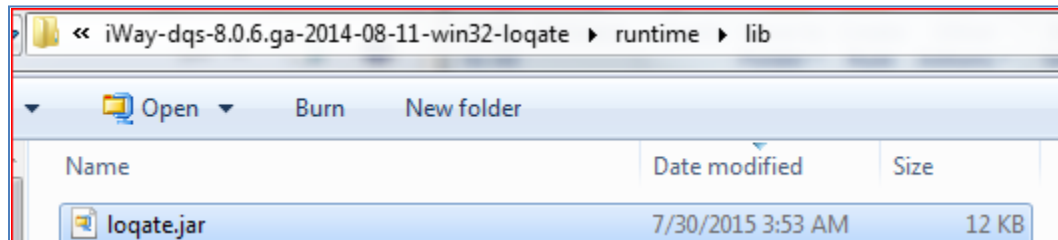


## Configuring the Data Quality Runtime

When Loqate is installed, it includes the latest Loqate API files will appear in your \loqate home directory.

The API files overwrite any previously installed API and must be copied into the DQ\runtime environment. This also applies to a DQ installed into iIT and the iSM server.

To configure the DQ runtime, copy the new *loqate.jar* file to the DQ runtime\lib folder, as shown in the following image.



As noted during the installation section, *lqtr.dll* and *lqtjava.dll* are already in the Loqate home directory, which has been automatically included in the system Path.

**Note:** You do not need to copy the *lqtr.dll* and *lqtjava.dll* files into the jre\bin, as long as this path remains accurate. Keep this in mind if you perform any updates.

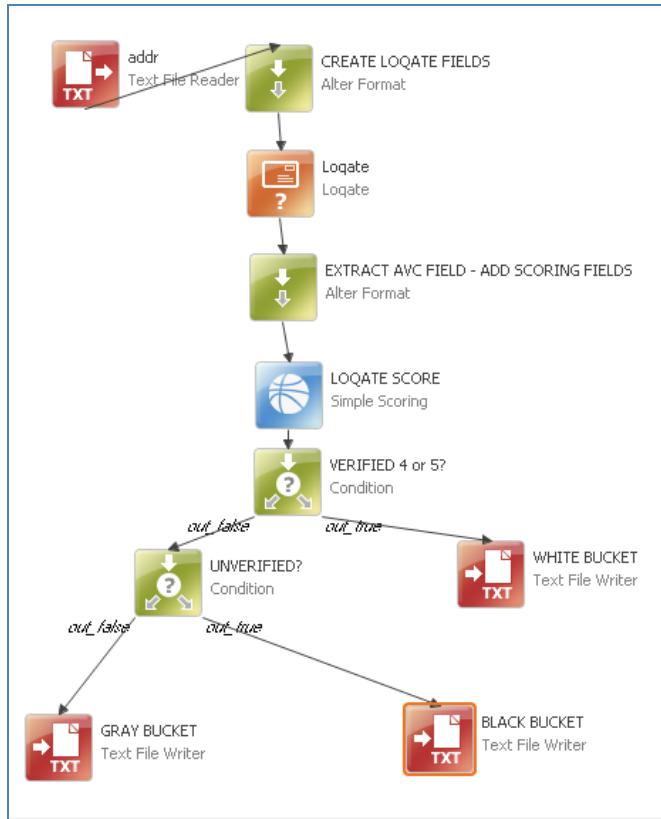
## Configuring the Data Quality Scoring Plan

Before you configure the Data Quality Scoring Plan, you must first:

1. Download and unzip the SCORING\_PLAN\_2016.zip file to any convenient directory.
2. Start your DQ GUI and navigate to the directory where you placed SCORING\_PLAN\_2016.zip and open SCORING\_PLAN.plan.

The DQ Scoring plan demonstrates how to set up a basic Loqate address checking plan. You only need to adjust the Loqate data folder setting to run the Loqate address check.

Address data records in CSV format, are read by the plan and then passed to the Loqate step, as shown in the following image.



The results are received in the form of an Address Verification Code (AVC) field.

In this application, this field is then parsed into separate columns for readability, and a score is assigned based on the quality of the address.

Each application you build will need to determine the appropriate level of what will constitute an acceptable address. For example, a *VERIFIED* address of *V5* is the highest possible score, while *UNVERIFIED* such as *U0* is unacceptable and should never be used. In between this range of scores is a vast gray area.

You will need to become familiar with the AVC codes in order to determine what is acceptable for your application.

To configure the plan, set the location of the Loqate data folder you installed, as shown in the following image.

▼ Other	
Data Folder*:	D:/LOQATE/2017Q1/data
Default Country:	US
Geocoding:	<input checked="" type="checkbox"/>
Input Elements*:	<a href="#">list element</a>
Output Elements*:	<a href="#">list element</a>

The input text file has about 76 records, as shown in the following sample image.

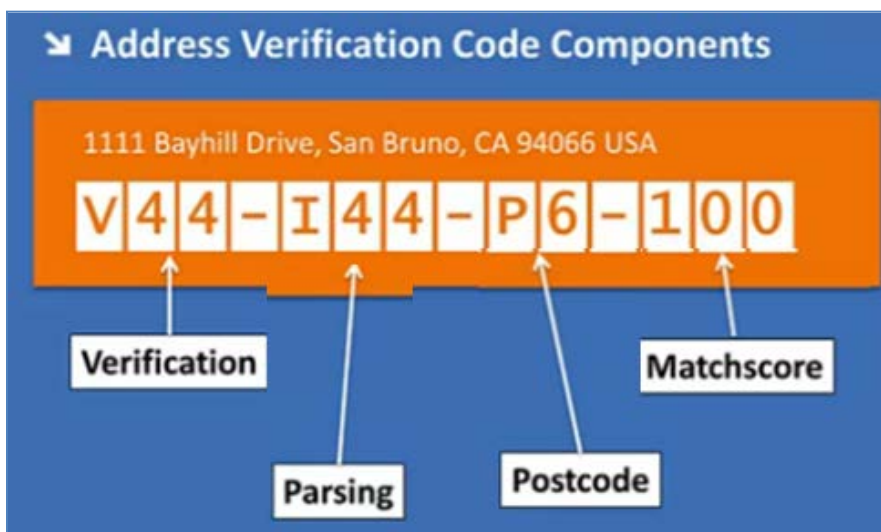
	Local_Address_1	Local_Address_2	Local_Address_3	Local_City	Local_County	Local_State	Local_Zip
1	NULL	P.O. Box 4229	NULL	Sacramento	SACRAMENTO	CA	95815
2	NULL	4790 Irvine Blvd., ...	NULL	Irvine	ORANGE	CA	92614
3	112 5 4th St. (98901)	or	Po Box 1789	Yakima	YAKIMA	WA	98907
4	NULL	PO BOX 11589	NULL	Spokane	SPOKANE	WA	99211
5	20 South Balph Ave.	NULL	NULL	Pittsburgh	ALLEGHENY	PA	15202
6	1036 WASHINGTON...	WILLIAMSPORT, ...	NULL	Williamsport	NULL	PA	17701
7	9320 PRIORITY WY ...	PO BOX 80883	INDIANAPOLIS IN	Indianapolis	NULL	IN	46240

The following images show how the input and output columns to the Loqate Element Names have been mapped.

Input Elements		Output Elements	
Column	Element Name	Column	Element Name
1 Local_Address_1	Address1	1 out_Address	Address
2 Local_Address_2	Address2	2 out_Address1	Address1
3 Local_Address_3	Address3	3 out_Address2	Address2
4 Local_City	Locality	4 out_Address3	Address3
5 Local_State	AdministrativeArea	5 AVC	AVC
6 Local_Zip	PostalCode	6 out_CITY	Locality
7 'US'	Country	7 out_STATE	AdministrativeArea
8 Local_County	SubAdministrativeArea	8 out_ZIP	PostalCode
*		9 out_LAT	Latitude
		10 out_LON	Longitude
		11 out_COUNTY	SubAdministrativeArea
		12 out_UNMATCHED	Unmatched
		13 out_Premise	Premise
		14 out_Thoroughfare	Thoroughfare

When Loqate analyzes an address, it returns a 14-character string known as the Address Verification Code (AVC). This plan shows you how to parse out the AVC and sort the results.

The following image shows a summary of how the AVC is formatted.



## Verifying Loqate Data

This section describes how to verify your Loqate data.

### Notes:

- The first character is the *Verification Status*, for example V = Verified
- The first digit represents the post-processing *Verification Level*. The first character and the first digit are the two most important factors.
- The second digit represents the pre-processing level in which it arrived.

The following tables list and describe the verification status values, characters, and levels used when verifying Loqate data.

Value	Status	Explanation
V	<i>Verified</i>	A complete match was made between the input data and a single GKR record. Values are 1 to 5
P	<i>Partially Verified</i>	Partial match made to a single GKR record.
U	<i>Unverified</i>	Unable to verify. The output fields will contain the input data.
A	<i>Ambiguous</i>	More than one close reference data match.
C	<i>Conflict</i>	More than one close reference data match with conflicting values.
R	<i>Reverted</i>	Record could not be matched to the specified minimum acceptable level (through Server Option: MinimumMatchscore). The output fields will contain the input data.

Value	Level	US
5	Delivery Point	Postbox or Suite
4	Premise	Premise or Building
3	Thoroughfare	Street
2	Locality	City
1	AdministrativeArea	State
0	None	

To make the results more understandable, the plan parses the AVC into separate fields. This is done with some simple string functions.

You can also apply optional scoring and explanations based on the AVC. This part is optional, though Scoring is used in most other parts of the DQ environment. In general, a higher score in DQ indicates a more serious defect. A score of 0 is perfect.

### Alter Format

General Added Columns Removed Columns

Added Columns:

	Name	Type	Expression
1	AVC_STATUS	STRING	substr(AVC,0,1 )
2	AVC_LEVEL	INTEGER	toInteger(substr(AVC,1,1 ))
3	AVC_MATCH_PCT	INTEGER	toInteger(substr(AVC,11,3 ))
4	AVC_ZIP	INTEGER	toInteger(substr(AVC,9,1 ))

To separate the good addresses, a filter is used. There are many ways you could do this. You have the option of using either the SCORE or using the AVC values that were parsed, as shown in the following image.

### Simple Scoring

General Advanced

General

Id: LOQATE SCORE

Other

Default Score Column: sco\_avc

Default Explain Column: exp\_avc

Scoring Cases:

	Condition	Desc...	Explanation	Score
1	AVC_STATUS='V'		'VERIFIED'	0
2	AVC_STATUS='P'		'PARTIAL'	100
3	AVC_STATUS='A'		'AMBIGUOUS'	250
4	AVC_STATUS='R'		'REVERTED'	500
5	AVC_STATUS='C'		'CONFLICT'	500
6	AVC_STATUS='U'		'UNVERIFIED'	1000
7	AVC_LEVEL = 0		"LVL_NONE"	2000
8	AVC_LEVEL = 1		"LVL_STATE"	1000
9	AVC_LEVEL = 2		"LVL_CITY"	500
10	AVC_LEVEL = 3		"LVL_STREET"	100
11	AVC_LEVEL = 4		"LVL_PREMISE"	10
12	AVC_LEVEL = 5		"LVL_DLV-PT"	0
13	AVC_MATCH_PCT < 95 and AVC...		"MATCH<95%"	50
14	AVC_MATCH_PCT < 85 and AVC...		"MATCH<85%"	100
15	AVC_MATCH_PCT < 75		"MATCH<75%"	200
16	AVC_ZIP >= 7		"ZIP_+4_OK"	0
17	AVC_ZIP >= 4 and AVC_ZIP < 7		"ZIP_PRIMARY_OK"	50

For example, V4 and V5 are considered good. This means the address contains a verified building/premise.

You can also add a check to ensure that the address is a close match. For example, if a minor typo has been corrected, then its match percentage might be slightly less than 100. However, if too much has changed for the address, then its match percentage is much lower. In this case, it may no longer represent the original value and you would not want to consider it as an acceptable (or “good”) address.

To prevent this situation, update the Condition to the following:

```
AVC_STATUS = 'V'
and AVC_LEVEL >= 4
and AVC_MATCH_PCT >= 85
```

As you become more familiar with the AVC codes, you can adjust any of these parameters. This will be based on how stringent your application needs to be, for example, are you delivering valuables or just sending flyers?

The records are split into three files. The following image shows a sample match percentage of 100.

out_Address1	out_Address2	out_Address3	out_CITY	out_ST...	out_ZIP	AVC	AVC_STATUS	AVC_LEVEL	AVC_MATCH_PCT
112 S 4th St 989...	PO Box 1789	Yakima WA 98...	Yakima	WA	98907-1789	V55-I55-P7-100	V	5	100
Null	PO Box 11589	Spokane WA 9...	Spokane	WA	99211-1589	V55-I55-P7-100	V	5	100
20 S Balph Ave	Pittsburgh PA...		Pittsburgh	PA	15202-3504	V44-I44-P7-100	V	4	100
2680 Grand Isla...	PO Box 308	Grand Island ...	Grand I...	NY	14072	V44-I55-P6-100	V	4	100
523 W 2nd St	Defiance OH ...		Defiance	OH	43512-2143	V44-I44-P7-100	V	4	100

Addresses with Unverified (U) symbols are considered bad and are written to a separate file, as shown in the following image.

out_Address	out_Address1	...	...	out...	..	...	AVC	AVC...	AVC_LEVEL	AVC_...	AVC_ZIP	sco_avc
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100
Null	Null			Null			U00-I02-P0-100	U	0	100	0	3100

Addresses containing the U symbol are not usable.

The remaining addresses are put into a third file. This represents that there is insufficient data and that it may require some analysis to be sent out or a remediation process to occur. This may also mean that human intervention may be required to make that determination.

### **Checking the Expiration Status for Loqate Data**

DQ Loqate data expires every year.

CASS data expires quarterly.

To check the expiration status on a Windows Operating System, run the following command:

```
lqtbatch -d [locate installation directory]\data
```

To check the expiration status on Unix, run the following command:

`lqbatch -d [locate installation directory]/data`

Capture the results and look for ExpiryDate, for example:

`DSEUSAC.ExpiryDate : 2017-12-31`

Each must be refreshed by downloading an update.

### Additional Resources

All users can create a support account with Loqate to gain access to the latest documentation by going to the following Loqate website:

<http://www.Loqate.com>

