



Installing and configuring R

Legal notice

Copyright

© 2017 LAVASTORM ANALYTICS, INC. ALL RIGHTS RESERVED. THIS DOCUMENT OR PARTS HEREOF MAY NOT BE REPRODUCED OR DISTRIBUTED IN ANY FORM WITHOUT THE WRITTEN PERMISSION OF LAVASTORM ANALYTICS, INC.

Apache Hive, Hive are trademarks of The Apache Software Foundation.

Apache Spark, Spark, Apache, and the Spark logo are trademarks of The Apache Software Foundation.

Microsoft and SharePoint are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

MongoDB and Mongo are registered trademarks of MongoDB, Inc.

Qlik®, Qlik Tech®, QlikView® and the Qlik Tech logos are trademarks or registered trademarks of Qlik Tech International AB.

Salesforce, SALESFORCE.COM and others are trademarks of salesforce.com, inc. and are used here with permission.

Tableau and Tableau logo are registered trademarks of Tableau Software, Inc.

TIBCO® Enterprise Runtime for R are either registered trademarks or trademarks of TIBCO Software Inc. and/or its subsidiaries in the United States and/or other countries.

The names of actual companies and products mentioned herein may be trademarks or registered trademarks of their respective owners.

Disclaimer

No representation, warranty or understanding is made or given by this document or the information contained within it, and no representation is made that the information contained in this document is complete, up to date or accurate. In no event shall LAVASTORM ANALYTICS, INC. be liable for damages or losses of any kind in connection with, or arising from its use, whether or not LAVASTORM ANALYTICS, INC. was made aware of the possibility of such damage or loss.



Tip: We recommend that you check our website for the latest documentation as minor updates or improvements may be made to the Help between releases.

Contact us

If you encounter any technical issues, we recommend that you visit the Dataverse Forums at <https://help.lavastorm.com/>. If your query has not been discussed previously in the forums, you can create a new topic and receive answers from our Dataverse experts. Alternatively, you can log a ticket by emailing support at help@lavastorm.com.

Our product is constantly evolving and input from you is highly valued. If you have any suggestions, please contact the product team at product@lavastorm.com.

Download

Lavastorm recommends that you use the latest version of the product. To download Dataverse, please go to <http://www.lavastorm.com/product-downloads/>.

Table of contents

Installing and configuring R	1
Installing and configuring R in Windows	1
Installing and configuring R in Linux	4

Installing and configuring R

Note: R and the RServe client are not distributed by Lavastorm. The R node communicates with these programs to execute R scripts.

Disclaimer



Open-source R is available under separate open source software license terms and is not part of Dataverse. As such, open-source R is not within the scope of your license for Dataverse. Open-source R is not supported, maintained, or warranted in any way by Lavastorm. Download and use of open-source R is solely at your own discretion and subject to the free open source license terms applicable to open-source R.

Installing and configuring R in Windows

The following steps are appropriate for the majority of Windows installations, and assume that the Dataverse server and R are on the same Windows machine. If extensive resources will be needed to process data using R scripts, R can be located on a separate server, if this is the case, please skip to [step 4](#).



Caution: If R is to be deployed on the same machine as Dataverse it is important to ensure sufficient system resources (especially RAM) are available to handle the peak processing load.

1. Install R as per the instructions on <https://www.r-project.org/>

The current version of the R node was tested against R 3.3.2-3.el7.x86_64.

2. Start R by using the Start menu created by the R installer.

An R window opens.

3. Install the additional packages required for the R node to execute by running the following commands inside your R window:

- `install.packages("Rserve")`
- `install.packages("jpeg")`
- `install.packages("png")`

The current version of the R node was tested against the following package versions (later versions of these libraries may not be compatible with the R node):

- Rserve 1.7-3
- png 0.1-7
- jpeg 0.1-8

4. Install the Rserve client into your existing Dataverse server, as follows:

- a. Navigate to Rserve's website (currently <http://rforge.net/Rserve/>) and download the Rserve_1.8-0.zip binary.



Note: Later versions of Rserve may not be compatible with the R node.

- b. Unzip the downloaded file, then navigate to Rserve/java. Copy the REngine.jar and Rserve.jar files.

- c. If you are using the server edition, paste the **REngine.jar** and **Rserve.jar** files into the following location: <Dataverse site directory>/lib/java

If you are using the desktop edition, paste the **REngine.jar** and **Rserve.jar** files into the following location:

%LOCALAPPDATA%/Dataverse/site/lib/java if you selected **Local User** during installation of Dataverse or C:/ProgramData/Dataverse/site/lib/java if you selected **All Users** during installation.

5. Download the R node LNA from the Lavastorm product downloads web page:

<http://www.lavastorm.com/product-downloads/>

Open Dataverse and import the R node LNA. See the integrated product help for more information on importing library nodes.

6. By default, the R node is imported into your **My Documents** folder.

If you are using the server edition, you can make the R node available for all users on the server by moving it to the **Public Documents** folder. To do this, from the Directory, click **My Documents**, then select the R node. Click the menu button in the details panel and select **Move to Public Documents**.



Caution: Prior to running the R node, you first need to [start the Rserve server](#). The Rserve server needs to be started whenever the machine on which it is running is rebooted.

Starting Rserve server

1. If you do not have an R window open, start R, using the Start menu icon created by the R installer.
2. Run the following commands in the R window:
 - `library(Rserve)`
 - `Rserve()`

You have now started the Rserve server. You can close the R window and the Rserve server will continue to run normally.



Note: There is no Startup menu item to shut down the Rserve server. To close Rserve, open Windows **Task Manager** and end the process named `Rserve.exe`.

Once you have completed the above steps, see the R node topic in the integrated product help for information on running the R node.

Installing and configuring R in Linux

The following steps assume that the Dataverse Server and R are on two separate machines. However, it is also possible to install them on the same machine.



Caution: If R is to be deployed on the same machine as Dataverse it is important to ensure sufficient system resources (especially RAM) are available to handle the peak processing load.

1. Install R as per the instructions on <https://www.r-project.org/>

The current version of the R node was tested against R 3.3.2-3.el7.x86_64.

If you choose to compile R from source, configure it with the following options to ensure that the supporting binaries for the R node are available:

- `--enable-R-shlib`
- `--with-libpng`
- `--with-jpeglib`

2. Start R by navigating to the R installation `bin` directory and executing the **R** program.

An R prompt opens, allowing you to enter R commands.

3. Install the additional packages that the R node needs to execute by typing the following commands in the R prompt window:

- `install.packages("Rserve")`
- `install.packages("jpeg")`
- `install.packages("png")`

The current version of the R node was tested against the following package versions (later versions of these libraries may not be compatible with the R node):

- Rserve 1.7-3
- png 0.1-7
- jpeg 0.1-8



Note: As CRAN packages for Linux are typically distributed as source packages, the Linux server must be provisioned with system libraries that permit the compilation of the prerequisite CRAN packages, e.g. gcc with the libraries for Fortran compilation.

4. Exit R by typing the command `q()`.
5. Start the X server on your machine, and ensure the identity that runs R has access to X. For most systems, this will mean running `xinit` as root and then ensuring that the `DISPLAY` environment variable is set to `:0` for the identity that starts Rserve. The identity that starts Rserve does not need to be in a graphical environment.
6. Confirm the location of the R temporary directory. The R node will use this directory to store image files as they are in transit from the R server to the Dataverse server. While the size of each image should be relatively small, they can add up over time, especially on systems with lots of users. The temporary directory can be set up via the standard R mechanisms (see R's `tempdir()` help for more information).
7. Configure Rserve as per the Rserve instructions at: <http://rforge.net/Rserve/doc.html#conf>. This is a three-step process:
 - Create a custom `Rserv.conf` configuration file.
 - Consider the security issues discussed in the section that begins "A note about security". The recommended security steps are important to help prevent unauthorized users from executing arbitrary R code on your R Server.
 - Ensure that the following options are in your `Rserv.conf` file:
 - `encoding utf8`
 - `interactive no`
 - `remote enable`
8. Navigate to the R bin directory and start the Rserve server by running the following command, where `<Rserv.conf>` is a reference to your custom `<Rserv.conf>` file:


```
R CMD Rserve --RS-conf <Rserv.conf>
```
9. Install the Rserve client on the Dataverse Server as per the following steps:
 - a. Download the `Rserve_1.8-3.tar.gz` file from the Rserve website: <http://rforge.net/Rserve/>.



Note: Later versions of Rserve may not be compatible with the R node.

- b. Unzip and untar the downloaded file, then navigate to: `Rserve/inst/java`.

- c. Add the **REngine.jar** and **Rserve.jar** files to the following location:

<Dataverse site directory>/lib/java

10. Download the R node LNA from the Lavastorm product downloads web page:

<http://www.lavastorm.com/product-downloads/>

Open Dataverse and import the R node LNA. See the integrated product help for more information on importing library nodes.

11. By default, the R node is imported into your **My Documents** folder.

If you are using the server edition, you can make the R node available for all users on the server by moving it to the **Public Documents** folder. From the Directory, click **My Documents**, then select the R node. Click the menu button in the details panel and select **Move to Public Documents**.

12. If you configured R to require authentication, ensure that your R node users are given their connection information to enable them to sign in. If the Dataverse and R Servers are on different machines, you should provide users with a hostname, and a port if you used a custom port for Rserve.

Once you have completed the above steps, see the R node topic in the integrated product help for information on running the R node.



© 2017 LAVASTORM ANALYTICS, INC.

www.lavastorm.com

Contact support: help@lavastorm.com

Document ID: DV-R-4

Date of publication: Monday, November 6, 2017