KNIME Interactive R Statistics Integration Installation Guide

KNIME AG, Zurich, Switzerland

Version 4.3 (last updated on 2020-12-06)
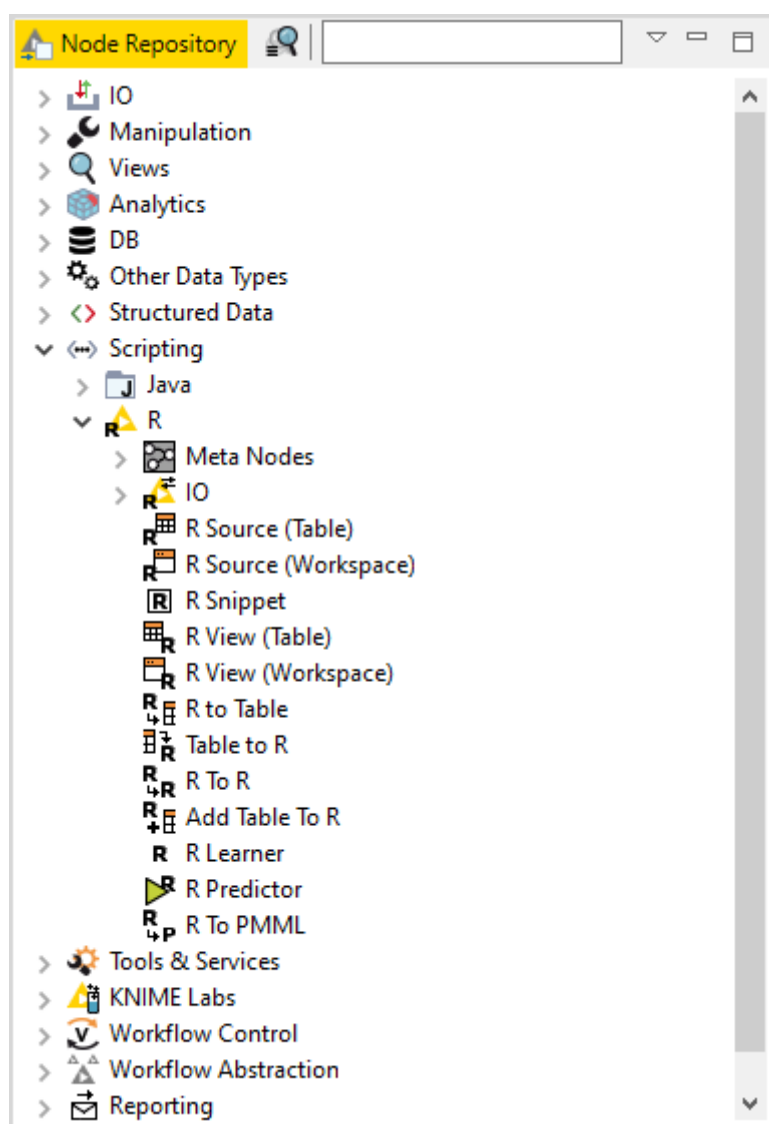
Table of Contents

Introduction

This guide describes how to install the KNIME Interactive R Statistics Integration to be used with KNIME Analytics Platform.

The KNIME Interactive R Statistics Integration allows to write and execute R scripts by interacting with an external R™ installation.

This integration comprises a handful of nodes which can be used to write R scripts and execute them in KNIME Analytics Platform. Besides providing the ability to read data from R into KNIME Analytics Platform, these nodes allow to manipulate data and create views using R, as well as to learn and apply models, trained in R, to the data.

These nodes can be found in the node repository under *Scripting → R*.

This integration, in order to communicate with R, depends on additional R packages which need to be installed. In this guide we describe how to configure the KNIME Interactive R

Statistics Integration, as well as how to install R and the necessary packages.

KNIME Interactive R Statistics Integration installation

The following steps are required to install and use the KNIME Interactive R Statistics Integration.

1. First, install the KNIME Interactive R Statistics Integration.

    In KNIME Analytics Platform, go to *File → Install KNIME Extensions*…. The KNIME Interactive R Statistics Integration can be found under *KNIME & Extensions* or by entering *R integration* into the search box.

2. Download and install R and all required packages as described in the R installation and R packages installation Sections.

3. Finally, configure the KNIME Interactive R Statistics Integration.

    In KNIME Analytics Platform go to *File → Preferences*. From the list on the left, select *R* under *KNIME* and the following dialog opens:
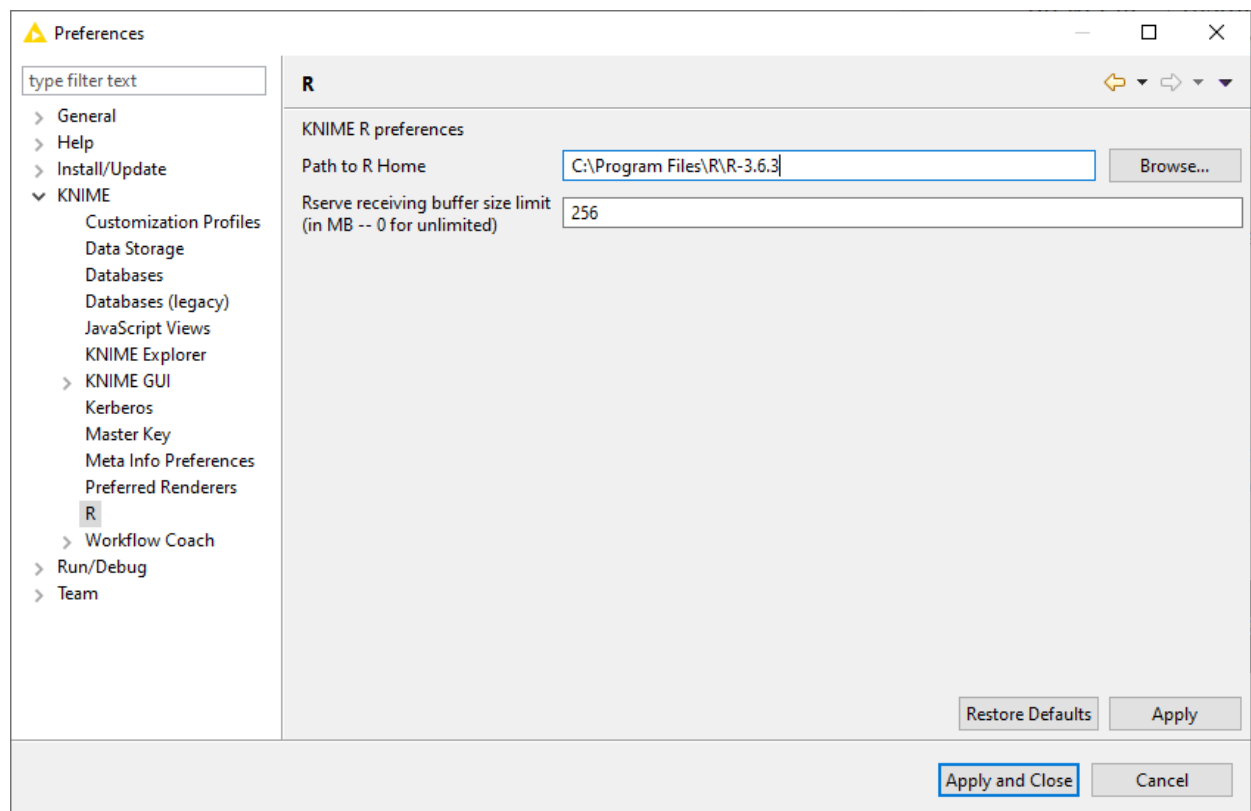
In *Path to R Home* enter the path to identify the location where R is installed and click

Apply to make the changes effective. This path can be found from within R by typing
`R.home()`.

R installation

The installation files as well as all the necessary information about R installation can be found on one of the mirror sites of the Comprehensive R Archive Network (CRAN). R installation can be completed safely using the default settings. The destination location where R will be installed will have to be specified in KNIME Analytics Platform and can be found from within `R` by typing `R.home()`. The 4.0.x versions of R are also supported.

ℹ
- For R installation on Windows the default location is recommended (e.g. usually `C:\Program Files\R\R-<version>`).

- On macOS systems the R installation folder location is usually `/Library/Frameworks/R.framework/Versions/Current/Resources/`.

- On Linux, instead, the installation folder is usually located in `/usr/lib/R`.

ℹ We recommend using R via the command-line interface and to avoid using `RStudio` or other tools as this might lead to unexpected behavior.

R packages installation

`Rserve` is an R package that allows other applications to talk to R using TCP/IP or local sockets sending binary requests. It requires different installation steps and different packages depending on the operating system in use.

Windows

1. Proceed with the installation of `Rserve`.

   From within R run the command:

   ```
   install.packages('Rserve')
   ```

   This will install the version 1.7-3.1 of `Rserve`.

   If the library folder is not writable a warning message will appear, together with the request to use a personal library:

   ```
   Warning in install.packages("Rserve") :
      'lib = "C:/Program Files/R/R-3.6.3/library"' is not writable
   Would you like to use a personal library instead? (yes/No/cancel) yes
   ```

   answer "yes" and a personal library will be installed and used.

   > **i** Please test the installation with KNIME Analytics Platform and in case a problem is encountered follow the instructions at point 2.

2. If the procedure at point 1 did not work properly, the installation of the latest version of `Rserve` might be necessary.

   a. First, from within `R`, remove any previously installed `Rserve` version running the command:

   ```
   remove.packages("Rserve")
   ```

   b. In this case, also `Rtools`, a collection of tools necessary for building R packages on Windows, needs to be installed. The installer is available here. In this guide `Rtools` is assumed to be installed under the default destination (e.g. usually `C:\Rtools\`).
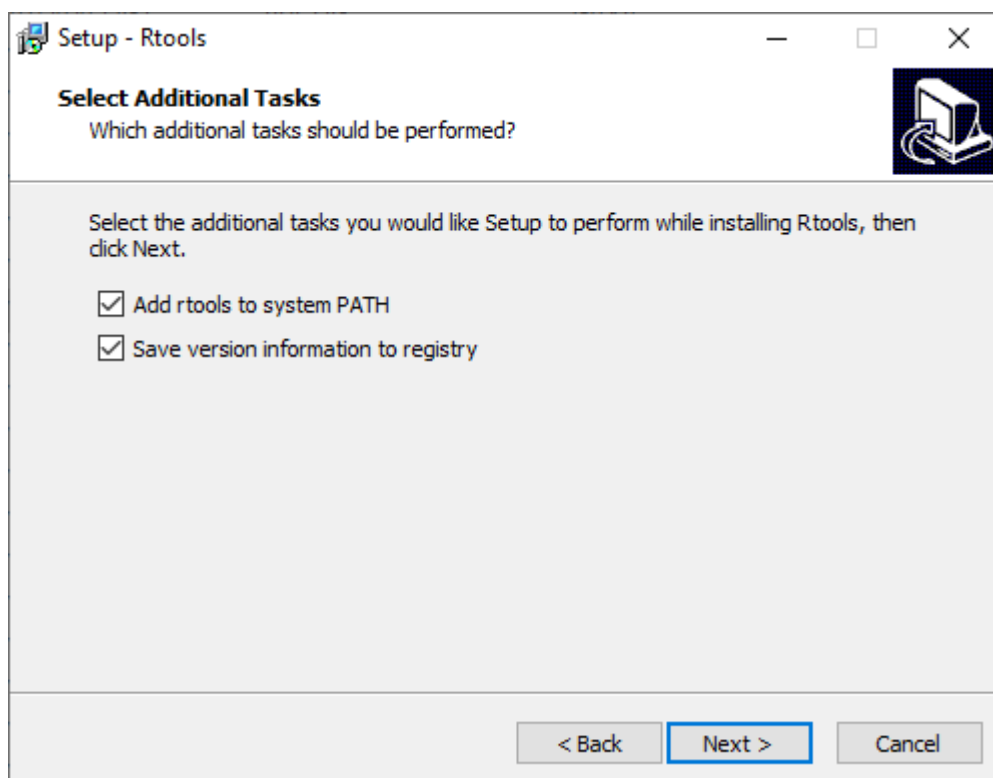
Check the box *Add rtools to system PATH* as shown in the above picture.

c. Please make sure that both R bin path and Rtools path are added to the Windows Path variable. From the *Start* menu open the *Control Panel* and click *System*. In the dialog that opens, click *Advanced system settings* in the left column. This opens the *System Properties* window. Here, click *Environment Variables…*. Then in the *System variables* section click *Edit…* to inspect and, if required, change the variable *Path*. Add the paths to R and Rtools bin folders, if they are not yet listed.
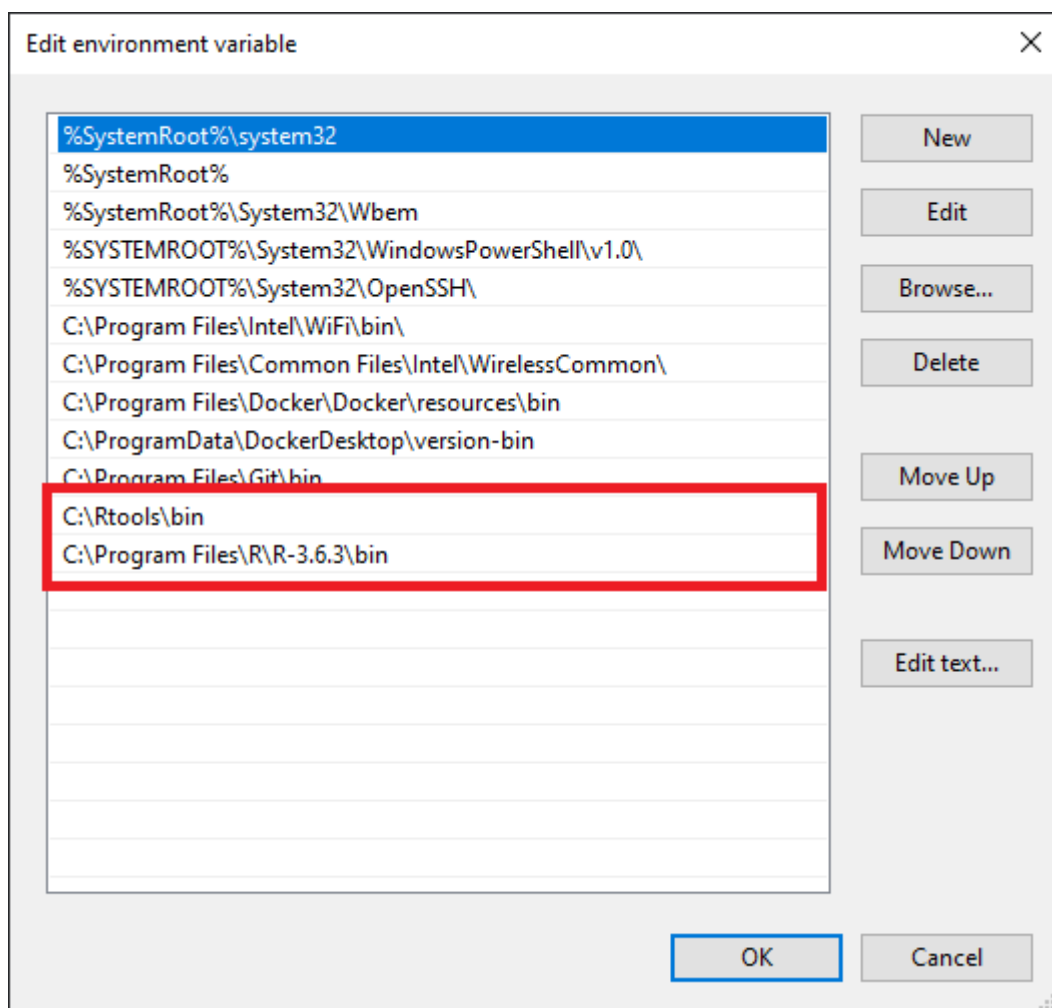
> ℹ️ For the latest version Rtools40 the `bin` folder is located in
> `C:\Rtools40\usr\bin`.

d. Finally, download and install the latest version of `Rserve`. From within R type the command:

```
install.packages('Rserve',,"http://rforge.net/",type="source")
```

MacOS

1. First, `Clang` and `GNU Fortran` are needed. If they are not already installed, the installation packages can be downloaded from this link.

2. Another program that needs to be installed is `XQuartz`, which is available for download at this link.

> ℹ️ A script is also available that can be used to install all of the necessary packages for R integration in KNIME Analytics Platform with macOS >= 10.14. After installing the KNIME Interactive R Statistics Integration the script can be found in `KNIME/plugins/org.knime.r_<knime_version>/scripts/` or at this link, only for KNIME Analytics Platform version 4.2 or higher. To allow for the execution of the script, please open a new *Terminal* and from the folder containing the script run `chmod u+x <name_of_the_script>.sh`.

3. Then the R package `Rserve` can be installed. Run `R` from *Terminal* and type the command:

```
install.packages('Rserve',,"http://rforge.net/",type="source")
```

The following error might appear when installing `Rserve`:

```
*** Rserve requires R (shared or static) library.              ***
*** Please install R library or compile R with either --enable-R-shlib ***
```

If your macOS is < 10.14 (Mojave), open a new *Terminal* window and type the following command:

```
xcode-select --install
```

otherwise please make use of the previously mentioned script.

Now continue with installing the `Rserve` package.

4. Finally, the R package `Cairo` needs to be installed. From within R type the command:

```
install.packages('Cairo')
```

Linux Ubuntu

To install `Rserve`, run `R` from *Terminal* and type the command:

```
install.packages('Rserve',,"http://rforge.net/",type="source")
```

If the following error message appears:

```
/usr/bin/ld: cannot find -lssl
collect2: error: ld returned 1 exit status
```

open a new *Terminal* window, type the following command:

```
sudo apt install libssl-dev
```

and continue with installing `Rserve`.

KNIME AG
Talacker 50
8001 Zurich, Switzerland
www.knime.com
info@knime.com