

Changes in R Version 2.8.0

by the R Core Team

Significant user-visible changes

- `var()`, `cov()`, `cor()`, `sd()` etc. now by default (when 'use' is not specified) return NA in many cases where they signalled an error before.

New features

- `abbreviate()` gains an optional argument 'strict' allowing cheap and fast strict abbreviation.
- The "lm" methods of `add1()`, `anova()` and `drop1()` warn if they are mis-used on an essentially exact fit.
- `as.array()` is now generic, gains a '...' argument.
- New function `as.hexmode()` for converting integers in hex format. `format.hexmode()` and `as.character.hexmode()` gain an 'upper.case' argument.
- `bitmap()` and `dev2bitmap()` gain support for anti-aliasing. The default type has been changed to 'png16m', which supports anti-aliasing.
- `Box.test()` gains a 'fitdf' argument to adjust the degrees of freedom if applied to residuals.
- `browseURL()` has a new argument 'encodeIfNeeded' to use `URLencode()` in cases where it seems likely that would be helpful. (Unfortunately, those are hard to guess.)
- `by()` gains a 'simplify' argument, passed to `tapply()`.
- `capabilities()` gains a new argument "tiff" to report if `tiff()` is operational.
- `chol2inv()` now treats `<numeric(1)>` as a [1 x 1]-matrix.
- `cov()` and `cor()` have the option 'use = "everything"' as default, and so does `var()` with its default 'na.rm = FALSE'. This returns NA instead of signalling an error for NA observations. Another new option is 'use = "na.or.complete"' which is the default for `var(*, na.rm=FALSE)`. `var(double(0), na.rm = L)` now returns NA instead of signalling an error, for both L = TRUE or FALSE, as one consequence of these changes.

- `data.matrix()` now tries harder to convert non-numeric columns, via `as.numeric()` or `as(, "numeric")`.
- `dev.interactive()` is able to recognize the standard screen devices if `getOption("device")` is a function (as well as by name).
- `dev.new()` gains a '...' argument which can be used to pass named arguments which will be used if appropriate to the device selected.
- `dimnames(x) <- value` extends 'value' if it is a list and too short, and 'x' is an array. This allows constructions such as `dimnames(x)[[1]] <- 1:3` to work whether or not 'x' already has dimnames.
- `format()`, `formatC()` and `prettyNum()` gain a new argument 'drop0trailing' which can be used to suppress trailing "0"s.
- `format()` now works for environments; also `print(env)` and `str(env)` share the same code for environments.
- It is now possible to create and open a text-mode `gzfile()` connection by explicitly using e.g. `open="rt"`.
- New `help.request()` function for compiling an e-mail to R-help according to "the rules". It is built on the new utility, `create.post()` on which also `bug.report()` is based now; both thanks to a contribution by Heather Turner.
- `help.search()` now assumes that non-ASCII items are in latin1 if that makes sense (all known examples on CRAN are).
- `HoltWinters()` and `decompose()` use a (statistically) more efficient computation for seasonal fits (they used to waste one period).
- `intToUtf8()` and `intToBits()` now accept numeric vectors, truncating them to integers.
- `is.unsorted()` gains an argument 'strictly'. It now works for classed objects with a `>=` or `>` method (as incorrectly documented earlier).
- `library()` no longer warns about masking objects that are identical(.,.) to those they mask.
- `lockBinding()`, `unlockBinding()`, `lockEnvironment()` and `makeActiveBinding()` now all return invisibly (they always return NULL).
- `mood.test()` now behaves better in the presence of ties.

- `na.action()` now works on fits of classes "lm", "glm",
- `optim(..., method="SANN", ..., trace=TRUE)` is now customizable via the 'REPORT' control argument, thanks to code proposals by Thomas Petzoldt.
- The 'factory-fresh' defaults for options ("device") have been changed to refer to the devices as functions in the `grDevices` namespace and not as names. This makes it more likely that the incorrect (since R 2.5.0) assumption in packages that `get(getOption("device"))()` will work will catch users of those packages.

- `pch=16` now has no border (for consistency with 15, 17, 18) and hence is now different from `pch=19`.

- `pdf()` has new arguments 'useDingbats' (set this to FALSE for use with broken viewers) and 'colormodel'. It now only references the ZapfDingbats font if it is used (for small opaque circles).

The default PDF version is now 1.4, since viewers that do not accept that are now rare.

Different viewers were rendering consecutive `text()` calls on a `pdf()` device in different ways where translucency was involved. The PDF generated has been changed to force each call to be rendered separately (which is the way `xpdf` or `ghostscript` was rendering, but Acrobat was forming a transparency group), which is consistent with other graphics devices supporting semi-transparency.

- `plot.dendrogram()` has new arguments (`xlim`, `yylim`) which allows zooming into a hierarchical clustering dendrogram.
- `plot.histogram()` gains an 'ann' argument. (Wish from Ben Bolker.)
- `plot(<lm_obj>)` now warns when it omits points with leverage one from a plot.
- `Plotmath` now recognizes 'aleph' and 'nabla' (the Adobe Symbol 'gradient' glyph) as symbol names.
- `polyroot()` no longer has a maximum degree.
- The `alpha/alphamax` argument of the 'nls' and 'mle' `profile()` methods is used to compute confidence limits for univariate t-statistics rather than a confidence region for all the parameters (and not just those being profiled).
- `quantile.default()` allows 'probs' to stray just beyond [0, 1], to allow for computed values.

- New functions `rawConnection()` and `rawConnectionValue()` allow raw vectors to be treated as connections.

- `read.dcf()` now consistently gives an error for malformed DCF.

- `read.fwf()` no longer passes its default for 'as.is' to `read.table()`: this allows the latter's default to be used.

- `readBin()`, `writeBin()`, `readChar()` and `writeChar()` now open a connection which was not already open in an appropriate binary mode rather than the default mode.

`readLines()`, `cat()` and `sink()` now open a connection which was not already open in an appropriate text mode rather than the default mode.

- `readCitationFile()` (and hence `citation`) now reads a package's CITATION file in the package's declared encoding (if there is one).

- The behaviour of `readLines()` for incomplete final lines on binary-mode connections has been changed to be like blocking rather than non-blocking text-mode connections.

- A new `reorder.character()` method has been added. This allows use of 'reorder(x, ...)' as a shorthand for 'reorder(factor(x), ...)' when 'x' is a character vector.

- `round()` now computes in long doubles where possible so the results are more likely to be correct to representation error.

- `rug()` now uses `axis()`'s new arguments from 2.7.2, hence no longer draws an axis line.

- `save()` (optionally, but by default) checks for the existence of objects before opening the file/connections (wish of PR#12543).

- `segments()`, `arrows()` and `rect()` allow zero-length coordinates. (Wish of PR#11192.)

- `set.seed(kind=NULL)` now takes 'kind' from a saved seed if the workspace has been restored or `.Random.seed` has been set in some other way. Previously it took the 'currently used' value, which was "default" unless random numbers had been used in the current session. Similarly for the values reported by `RNGkind()`. (Related to PR#12567.)

`set.seed()` gains a 'normal.kind' argument.

- `setEPS()` and `setPS()` gain '...' to allow other arguments to be passed to `ps.options()`, including overriding 'width' and 'height'.

- `setTimeLimit()` function to set limits on the CPU and/or elapsed time for each top-level computation, and `setSessionLimit()` to set limits for the rest of the session.
- `splinefun()` has a new method = "monoH.FC" for monotone Hermite spline interpolation.
- `sprintf()` optionally supports the %a/%A notation of C99 (if the platform does, including under Windows).
- `str()`'s default method gains a 'formatNum' function argument which is used for formatting numeric vectors. Note that this is very slightly not backward compatible, and that its default may change before release.
- The `summary()` method for class "ecdf" now uses a `print()` method rather than printing directly.
- `summary.manova()` uses a stabler computation of the test statistics, and gains a 'tol' argument to allow highly correlated responses to be explored (with probable loss of accuracy). Similar changes have been made to `anova.mlm()` and `anova.mlmlist()`.
- `Sweave()` now writes concordance information inside a `\Sconcordance` \LaTeX macro, which allows it to be inserted into PDF output.
- `system.time()` now uses lazy evaluation rather than `eval/substitute`, which results in more natural scoping. (PR#11169.)
- In `table()`, 'exclude=NULL' now does something also for factor arguments. A new 'useNA' argument allows you to control whether to add NA levels unconditionally or only when present in data. A new convenience function `addNA()` gives similar functionality by adding NA levels to individual factors.
- `unlink()` tries the literal pattern if it does not match with wildcards interpreted – this helps with e.g. `unlink("a[b")` which previously needed to be `unlink("a\\[b")`.
- `update.packages()` gains an argument 'oldPkgs', where `new.packages()` and `old.packages()` get 'instPkgs'. These allow to consider only subsets of packages instead of all installed ones.
- `which(b)` is somewhat faster now, notably for named vectors, thanks to a suggestion by Henrik Bengtsson.
- New generic function `xtfrm()` as an auxiliary helper for `sort()`, `order()` and `rank()`. This should return a numeric vector that sorts in the same way as its input. The default method supports any class with `==`, `>` and `is.na()` methods but specific methods can be much faster. As a side-effect, `rank()` will now work better on classed objects, although possibly rather slowly.
- `X11()` and `capabilities("X11")` now catch some X11 I/O errors that previously terminated R. These were rare and have only been seen with a misconfigured X11 setup on some versions of X11.
- The handling of nuls in character strings has been changed – they are no longer allowed, and attempting to create such a string now gives a truncation warning (unless `options("warnEscapes")` is FALSE).
- The user environment and profile files can now be specified via environment variables 'R_ENVIRON_USER' and 'R_PROFILE_USER', respectively.
- `?pkg::topic` and `?pkg:::topic` now find help on 'topic' from package 'pkg' (and not help on :: or :::).
- `??topic` now does `help.search("topic")`; variations such as `??pkg::topic` or `field??topic` are also supported.
- There is support for using ICU (International Components for Unicode) for collation, enabled by configure option `--with-ICU` on a Unix-alike and by a setting in MkRules on Windows. Function `icuSetCollate()` allows the collation rules (including the locale) to be tuned. [Experimental.]
- If S4 method dispatch is on and S4 objects are found as attributes, `show()` rather than `print()` is used to print the S4 attributes.
- Starting package `tcltk` without access to Tk (e.g. no available display) is now a warning rather than an error, as Tcl will still be usable. (On most platforms it was possible to inhibit Tk by not having DISPLAY set, but not on Windows nor Mac OS X builds with `--with-aqua`.)
- Using `$` on a non-subsettable object (such as a function) is now an error (rather than returning NULL).
- Hexadecimal numerical constants (such as `0xab.cdp+12`) may now contain a decimal point.
- PCRE has been updated to version 7.8 (mainly bug fixes).
- `plot.ecdf()` now defaults to `pch=19` so as to better convey the left-closed line segments.

New features in package 'methods'

- S3 classes that are registered by a call to `setOldClass()` now have the S3 class as a special slot, and therefore so do any S4 classes that contain them. This mechanism is used to support S4 classes that extend S3 classes, to the extent possible. See `?Classes`, `?setOldClass`, and `?S3Class`.

The treatment of special pseudo-classes "matrix", "array", and "ts" as S4 classes has also been modified to be more consistent and, within limitations imposed by special treatment of these objects in the base code, to allow other classes to contain them. See `class?ts`.

A general feature added to implement "ts" and also "data.frame" as S4 classes is that an S4 class definition can be supplied to `setOldClass()` when the S3 class has known attributes of known class.

`setOldClass()` now saves all the S3 inheritance, allowing the calls to be built up in stages, rather than including all the S3 classes in each call. Also allows `as(x, "S3")` to generate valid S3 inheritance from the stored definition. See `?S3`.

- S4 methods may now be defined corresponding to "...", by creating a generic function that has "..." as its signature. A method will be selected and called if all the arguments matching "..." are from this class or a subclass. See `?dotsMethods`.
- New functions `S3Part()` and `S3Class()` provide access to the corresponding S3 object and class for S4 classes that extend either an S3 class or a basic R object type.
- `show(<class definition>)` now also shows the class name.

Installation

- If sub-architectures are used, a copy of `Rscript` is installed in `$R_HOME/bin/exec$R_ARCH` (since that in `$R_HOME/bin` and `/usr/bin` might be overwritten in a subsequent installation).

Package installation

- `LazyLoad`: yes is now the default, so packages wanting to avoid lazy loading must set 'LazyLoad: no' (or an equivalent value) in the `DESCRIPTION` file.

- R CMD INSTALL will now fail if it finds a non-executable 'configure' script in the package – this usually indicates a file system with insufficient permissions. If a non-executable 'cleanup' script is found and either `-clean` or `-preclean` is used, a warning is given.

Deprecated & defunct

- Use in packages of the graphics headers `Rdevices.h` and `Rgraphics.h` is defunct: they are no longer installed.
- `options("par.ask.default")` is defunct in favour of `"device.ask.default"`.
- The 'device-independent' family "symbol" is defunct: use `font=5` (base) or `fontface=5` (grid) instead.
- `gammaCody()` is defunct.
- `par("gamma")` is defunct.
- 'methods' package functions `getAccess()`, `getAllMethods()`, `getClassName()`, `getClassPackage()`, `getExtends()`, `getProperties()`, `getPrototype()`, `getSubclasses()`, `getVirtual()`, `mlistMetaName()`, `removeMethodsObject()` and `seemsS4Object()` are defunct.
- Use of a non-integer `.Random.seed` is now an error. (R itself has never generated such values, but user code has, and R >= 2.6.0 has given a warning.)
- `methods::allGenerics()` is deprecated.
- In package installation, `SaveImage: yes` is now ignored, and any use of the field will give a warning.
- `unserialize()` no longer accepts character strings as input.
- The C macro 'allocString' has been removed – use 'mkChar' and variants.
- Use of `allocVector(CHARSXP ...)` is deprecated and gives a warning.

Utilities

- The default for 'stylepath' in Sweave's (default) `RweaveLatex` driver is now `FALSE` rather than `TRUE` if `SWEAVE_STYLEPATH_DEFAULT` is unset: see `?RweaveLatex`. To support this, `tools::texi2dvi` adds the R 'texmf' directory to the input search path.

- R CMD Rd2dvi now previews PDF output (as was documented) if R_PDFVIEWER is set (as it will normally be on a Unix-alike but not on Windows, where the file association is used by default).
- R CMD check checks for binary executable files (which should not appear in a source package), using a suitable 'file' if available, else by name.
- R CMD check now also uses codetools' checks on the body of S4 methods.

C-level facilities

- R_ReadConsole will now be called with a buffer size of 4096 bytes (rather than 1024): maintainers of alternative front-ends should check that they do not have a smaller limit.
- Graphics structure NewDevDesc has been renamed to DevDesc. For now there is a compatibility define in GraphicsDevice.h, but it will be removed in R 2.9.0.
- PROTECT and UNPROTECT macros now work even with R_NO_REMAP.

Bug fixes

- @ now gives an error (and not just a warning) if it is being applied to a non-S4 object.
- R CMD appends (not prepends) R's texmf path to TEXINPUTS.
- Objects generated by new() from S4 classes should now all satisfy isS4(object). Previously, prototypes not of object type S4 would not be S4 objects. new() applied to basic, non-S4 classes still will (and should) return non-S4 objects.
- Functions writing to connections such as writeLines(), writeBin(), writeChar(), save(), dput() and dump() now check more carefully that the connections are opened for writing, including connections that they open themselves.
Similarly functions which read such as readLines(), scan(), dcf() and parse() check connections for being open for reading.
- Equality comparison of factors with <NA> levels now works correctly again.
- Repainting of open X11 View() windows is now done whilst an X11 dataentry window is in use.

- Indexing of data frames with NA column names and a numeric or logical column index works again even if columns with NA names are selected.
- on.exit() has been fixed to use lexical scope in determining where to evaluate the exit action when the on.exit expression appears in a function argument.
- rank() now consistently returns a double result for ties.method = "average" and an integer result otherwise. Previously the storage mode depended on 'na.last' and if any NAs were present.
- The "lm" methods of add1(), and drop1() now also work on a model fit with na.action = na.exclude.
- median(c(x = NA_real_)) no longer has spurious names().
- isoreg(x, y) now returns the correct result also when x has ties, in all cases.
- What na.action() does is now correctly documented.
- source() with echo=TRUE now behaves like ordinary automatic printing, by using methods::show() for S4 objects.
- Several bugs fixed in '?' with topics: it previously died trying to construct some error messages; for S4 methods, class "ANY" should be used for omitted arguments and default methods.
- trace() should create missing traceable classes in the global environment, not in baseenv() where other classes will not be found.
- Class inheritance using explicit coerce= methods via setIs() failed to coerce the argument in method dispatch. With this fixed, a mechanism was needed to prohibit such inheritance when it would break the generic function (e.g., initialize). See ?setIs and ?setGeneric.
- RSiteSearch() encodes its query (it seems this is occasionally needed on some platforms, but encoding other fields is harmful).
- 'incomparables' in match() was looking up indices in the wrong table.
- write.dcf() did not escape "." according to Debian policy (PR#12816).
- col2rgb() sometimes opened a graphics device unnecessarily, and col2rgb(NA) did not return a transparent color, as documented.

- `pdf(family="Japan")` [and other CIDfonts] no longer seg.faults when writing "western" text strings.
- `as.list()` applied to an environment now forces promises and returns values.
- Promises capturing calls to `sys.parent()` and friends did not work properly when evaluated via method dispatch for internal S3 generics.
- The default `pkgType` option for non-CRAN builds of R on Mac OS X is now correctly "source" as documented.
- The default `R_LIBS_USER` path in AQUA builds now matches the Mac-specific path used by the Mac GUI: `/Library/R/x.y/library`
- `splinefun()` with natural splines incorrectly evaluated derivatives to the left of the first knot. (PR#13132, fix thanks to Berwin Turlach.)
- `anova(glm(..., y=FALSE))` now works. (PR#13098.)
- `cut.Date(x, "weeks")` could fail if `x` has only one unique value which fell on a week boundary. (PR#13159.)