

Package ‘R1010’

April 1, 2015

Type Package

Title R interface to 1010data

Version 1.0

Date 2014-08-18

Author Chris Simon <chris.simon@1010data.com>, Jeffery Taylor <jeffery.taylor@1010data.com> (Documentation)

Maintainer Support <support@1010data.com>

Description A set of functions that help a 1010data user interact with 1010data through R.

License see LICENSE.txt

Depends Rcpp,pryr

Suggests testthat

LinkingTo Rcpp

Collate onAttach.R util.R valueExprParser.R newSession.R session.R openTable.R query.R colord.R getDir.R head.R tail.R info.R sel.R tabu.R willbe.R link.R

R topics documented:

R1010-package	2
clearCache	2
dim	3
getCurrMem	3
getDir	4
getSystemVersion	5
head.R1010	5
info	6
link	7
logoutSession	8
newSession	8
nrow	9
openTable	10
posSession	11
query	12

sel	13
tabu	14
tail.R1010	15
willbe	16

Index	17
--------------	-----------

R1010-package	<i>An interface between 1010data and R.</i>
---------------	---

Description

The functions in R1010 provide a 1010data user the ability to manage sessions, view tables, and run queries.

Details

Package:	R1010
Type:	Package
Version:	0.1.2
Date:	2014-09-24
License:	see LICENSE.txt

Author(s)

chris.simon@1010data.com

References

www.1010data.com

clearCache	<i>Clears server side cache.</i>
------------	----------------------------------

Description

Frees data stored in the cache on 1010data's server for the given session.

Usage

```
clearCache(session)
```

Arguments

session A session handle generated using newSession.

Examples

```
## Not run:
s <- posSession("captain_smith")
clearCache(s)

## End(Not run)
```

dim *Gets the number of rows and columns in a 1010data table.*

Description

Returns the nrows attribute and the length of the given data.frame.

Usage

```
dim(df)
```

Arguments

df An object of class R1010 and data.frame.

Examples

```
## Not run:
s <- newSession("captain_smith", "meLike1010")
df <- openTable(s, "pub.demo.retail.item")
dim(df) # [1] 35 8

## End(Not run)
```

getCurrMem *Get memory allocated to workspace.*

Description

Returns the sum of the memory high water marks on each machine that your ID has access to. This indicates how much total memory has been allocated to your session.

Usage

```
getCurrMem(session)
```

Arguments

session A session handle generated using newSession.

Examples

```
## Not run:
s <- newSession("captain_smith", "meLike1010")
memory.allocated <- getCurrMem(s)

## End(Not run)
```

getDir *Get information about a particular directory (folder) on 1010data.*

Description

Returns a data.frame of the tables and folders visible from the given folder path. The default call provides a data.frame of the current tables and folders visible from the root directory.

Usage

```
getDir(ses, path = "", recursive = FALSE)
```

Arguments

ses A 1010data session handle generated with newSession.

path A valid 1010data folder path.

recursive An object of type boolean that specifies whether to return directory information about all subdirectories. The default is FALSE.

Examples

```
## Not run:
s <- newSession("captain_smith", "meLike1010")
getDir(s) # folders at root
getDir(s, "pub") # folders and tables in pub folder
dir(s, "pub.demo") # folders and tables in pub.demo folder
getDir(s, "pub.demo", TRUE) # all folders under pub.demo recursively
getDir(s, "path.to.table") # error, as tables are not directories

## End(Not run)
```

getSystemVersion	Returns the 1010data system version.
------------------	--------------------------------------

Description

Returns the 1010data system version associated with a specific R1010 session handle. The system version is a string which details the most recent version that a handle successfully logged on to.

Usage

```
getSystemVersion(session)
```

Arguments

session	A session handle generated using newSession.
---------	--

Examples

```
## Not run:  
s <- newSession("captain_smith", "meLike1010")  
system.version <- getSystemVersion(s) # [1] "prod-6.99"  
  
## End(Not run)
```

head.R1010	View the first n rows of the current R1010 table object.
------------	--

Description

Allows the user to view the current state of a table (possibly transformed by a query).

Usage

```
head(df, nrows = 20)
```

Arguments

df	An object of class R1010 and data.frame.
nrows	The number of rows to view from the top of the table. The default is 20.

Examples

```
## Not run:
ses <- newSession("moe1010", "P13a5e")
df <- openTable(ses, "pub.demo.weather.stations")

dfq.select <- query(df,
  "<sel value=\"between(i_;5;21)\"/>")

head(df) # shows first 20 rows of pub.demo.weather.stations
head(dfq.select) # shows all 17 rows after <sel> operation
head(dfq.select, 10) # shows first 10 rows after <sel> operation

## End(Not run)
```

info

Returns information about the various attributes of a 1010data table.

Description

Returns the table path, column information, dimensions, current memory highwater mark, and system version associated with the given data.frame.

Usage

```
info(df)
```

Arguments

df An object of class R1010 and data.frame.

Examples

```
## Not run:
s <- newSession("captain_smith", "meLike1010")
df <- openTable(s, "pub.demo.retail.item")
info(df)

## End(Not run)
```

link	<i>Function to link a foreign table of class R1010 to a base table of class R1010.</i>
------	--

Description

Allows the linking of two R1010 tables together.

Usage

```
link(df1, df2, by, by2 = NULL, options = NULL, run = FALSE, ...)
```

Arguments

df	An object of class R1010 and data.frame that represents the base table.
df2	An object of class R1010 and data.frame that represents the foreign table. Operations on the foreign table are applied prior to effecting the link, so that the current table is linked to the foreign table as transformed by those operations.
by	An object of type character that represents the linking column in the base table.
by2	An object of type character that represents the linking column in the foreign table. The by2 parameter may be omitted if the linking column in the foreign table has the same name as the linking column in the base table.
options	An object of type list that represents valid link attributes. The default is NULL. See the 1010data Reference Manual for valid attributes. (Note: The attributes table2, col, and col2 should not be specified via the options argument; they should instead be specified via the df2, by, and by2 parameters.)
run	An object of type boolean that specifies whether the tables will be linked. If FALSE, the link is added to the query in the df argument, but the tables are not linked. The default is FALSE.

Details

The link function requires a table handle (an object of class R1010) for the base table, a table handle (an object of class R1010) for the foreign table, the name of a linking column in the base table, the name of a linking column in the foreign table (if different from the base table), and it returns a table handle containing the results of the link (if run is set to TRUE).

Examples

```
## Not run:  
ses <- newSession("curly1010", "P13a5e")  
  
df <- openTable(ses, "pub.demo.retail.item")  
df.prod <- openTable(ses, "pub.demo.retail.prod")  
df.linked <- link(df, df.prod, "sku", run=TRUE)
```

```
df.prod.sel <- sel(df.prod, "dept==22", run=TRUE)
df.linked.sel <- link(df, df.prod.sel, "sku", option=list(type="select", suffix="_prod"), run=TRUE)

# Example using pipeR
library(pipeR)
df %>>% link(df.prod.sel, "sku", option=list(type="select", suffix="_prod")) %>>% head
```

```
## End(Not run)
```

logoutSession	<i>Log out of 1010data session</i>
---------------	------------------------------------

Description

Used for logout event.

Usage

```
logoutSession(session)
```

Arguments

session A 1010data session handle generated with newSession.

newSession	<i>Generates a session handle for subsequent R1010 function calls.</i>
------------	--

Description

Returns a R1010 data session handle.

Usage

```
newSession(username, password = "", kill = "notKill",
            gateway = "http://www2.1010data.com/cgi-bin/gw.k",
            xml.log.list = list("off", ""), log.name = "")
```


Arguments

username	An object of type character that represents a valid 1010data username.
password	An object of type character that represents a valid 1010data password. If an empty string (e.g., "") is passed, the user is prompted to enter a password. The default is an empty string.
kill	An object of type character that represents a valid 1010data kill value. If kill = "pos", a live session is possessed or new session is created. If kill = "kill", current live session is killed, and a new session is created. If kill = "notKill", current live session if not killed and no new session is created, or if there is no current live session a new session is created. The default is kill = "notKill".
gateway	An object of type character that represents a valid 1010data gateway. The default is "http://www2.1010data.com/cgi-bin/gw.k".
xml.log.list	An object of type list with two elements that determines if XML logging should be enabled. The first element may be "off" or "on", and the second element is a filename.
log.name	If a non-empty object of type character is passed to log.name, then logging is enabled. This will not enable XML logging.

Details

A call to the newSession function produces a variable that points to the location in memory where session information is stored. It is used by other functions in the R1010 package to interact with a 1010data session. It also allows for two different forms of logging.

Examples

```
## Not run:
ses <- newSession("Moe1010", "P13A5e")
df <- openTable(ses, "pub.demo.weather.stations")

## End(Not run)
```

nrow

Get number of rows in a 1010data table.

Description

Returns the nrows attribute of the given data.frame.

Usage

```
nrow(df)
```

Arguments

df	An object of class R1010 and data.frame.
----	--

Examples

```
## Not run:
s <- newSession("captain_smith", "meLike1010")
df <- openTable(s, "pub.demo.retail.item")
retail.item.rows <- nrow(df)

## End(Not run)
```

openTable	<i>Attach a 1010data session handle to a table on 1010data and return a data.frame.</i>
-----------	---

Description

Used to open a table stored in 1010data and view as data.frame.

Usage

```
openTable(session, tab, row.range = NULL, header = "name")
```

Arguments

session	A session handle generated using newSession.
table	An object of type character that represents a path to a table on 1010data.
row.range	An object of type integer (or double) with length equal to 2. This object represent the range of rows to download. The first value specifies the starting row, and the second value specifies the last row. the row.range argument may also take the character values "none" or "all". "none" will return a valid table handle, but with an empty data.frame. "all" will attempt to download all rows.

Examples

```
## Not run:
s <- newSession("captain_smith", "meLike1010")
df.0 <- openTable(s, "pub.demo.weather.stations")
df.1 <- openTable(s, "pub.demo.weather.stations", row.range=c(1,45))
df.2 <- openTable(s, "uploads.t99956564297_captain_smith", row.range = "none")

# get character vector of all table paths in pub.demo.baseball
paths <- dir(s,"pub.demo.baseball")[,2]
# store table handles for all tables in pub.demo.baseball
baseball.tables <- lapply(paths,function(x) openTable(s,x))
# get number of rows for each table in pub.demo.baseball
sapply(baseball.tables,nrow)
# [1] 78881 15350 32896 59884 3707 2895 7439 2776
# get character vector of all column names in pub.demo.baseball
unique(unlist(sapply(baseball.tables,names)))

## End(Not run)
```

posSession	<i>Possess an existing 1010data session.</i>
------------	--

Description

Returns the session handle for an existing 1010data session.

Usage

```
posSession(username, password = "",  
           xml.log.list = list("off", ""), log.name = "")
```

Arguments

username	An object of type character that represents a valid 1010data username.
password	An object of type character that represents a valid 1010data password. If an empty string (e.g., "") is passed, the user is prompted to enter a password. The default is an empty string.
xml.log.list	An object of type list with two elements that determines if XML logging should be enabled. The first element may be "off" or "on", and the second element is a filename.
log.name	If a non-empty object of type character is passed to log.name, then logging is enabled. This will not enable XML logging.

Details

A call to the posSession function produces a variable that points to the location in memory where session information is stored. It is used by other functions in the R1010 package to interact with a 1010data session. It also allows for two different forms of logging.

Examples

```
## Not run:  
ses <- posSession("Moe1010")  
df <- openTable(ses, "pub.demo.weather.stations")  
  
## End(Not run)
```

query	<i>Function to query tables on 1010data.</i>
-------	--

Description

Allows the user to run a 1010data Macro Language query.

Usage

```
query(df, q, file = NULL, row.range = NULL, run = TRUE, colord = NULL, ...)
```

Arguments

df	An object of class R1010 and data.frame.
q	An object of type character that represents a query written in the 1010data Macro Language.
file	An object of type character that represents the name of a file in the working directory which contains a 1010data query.
row.range	An object of type integer (or double) with length equal to 2. This object represent the range of rows to download. The first value specifies the starting row, and the second value specifies the last row. the row.range argument may also take the character values "none" or "all". "none" will return a valid table handle, but with an empty data.frame. "all" will attempt to download all rows.
run	An object of type boolean that specifies whether the query is to be run. The default is TRUE.
	/itemcolord Maybe NULL or a character vector. If colord = NULL the current set of columns in view are returned, else a character vector of length 1 that delimits with commas or a character vector that has column names as elements.

Details

The query function may be run with a string containing the query, or a string that specifies a file-name that contains a valid 1010data query. A data.frame containing session information (attached using openTable) may also be used in the df argument with run=FALSE to prevent unnecessary API requests.

Examples

```
## Not run:
ses <- newSession("moe1010", "P13a5e")
df <- openTable(ses, "pub.demo.weather.stations")

dfq.select <- query(df,
  "<sel value=\"between(i_;5;21)\"/>")

dfq.select.and.willbe.1 <- query(dfq.select,
  "<willbe name=\"is_az\" value=\"state=AZ\"/>")
```

```

dfq.select.and.willbe.2 <- query(df,
  "<sel value=\"between(i_;5;21)\"/>"
  "<willbe name=\"is_az\" value=\"state=AZ\"/>" )

identical(dfq.select.and.willbe.1,dfq.select.and.willbe.2) # TRUE

dfq.select.and.willbe.3 <- query(
  query( df , "<sel value=\"between(i_;5;21)\"/>" run = FALSE ),
  "<willbe name=\"is_az\" value=\"state=AZ\"/>" )

identical(dfq.select.and.willbe.1,dfq.select.and.willbe.3) # TRUE

# Example use of colord
query(df,"<sel value=\"between(i_;5;21)\"/>",colord="id,state")
query(df,"<sel value=\"between(i_;5;21)\"/>",colord=c("id","state","lat","lon"))

## End(Not run)

```

sel

Function to perform a selection on a 1010data table.

Description

Allows the user to perform a selection on a 1010data table.

Usage

```
sel(df, expr, run = FALSE, ...)
```

Arguments

df	An object of class R1010 and data. frame.
expr	An object of type character that is a valid 1010data expression. See the 1010data Reference Manual .
run	An object of type boolean that specifies whether the selection is to be run. If FALSE, the selection is added to the query in the df argument, but the selection is not performed. The default is FALSE.

Details

The sel function requires a table handle (an object of class R1010) and an expression, and it returns a table handle containing the results of the selection (if run is set to TRUE).

Examples

```
## Not run:
ses <- newSession("moe1010", "P13a5e")
df <- openTable(ses, "pub.demo.retail.item")

sel(df, "account=957", run=TRUE)

sel(df, "between(date; 20120515; 20120516)", run=T)

sel(df, "g_first1(store date; ; )", run=T)

## End(Not run)
```

tabu	<i>Function to perform a tabulation or cross-tabulation on a 1010data table.</i>
------	--

Description

Allows the user to perform a tabulation on a 1010data table.

Usage

```
tabu(df, breaks, tcols = NULL, cbreaks = NULL, run = FALSE)
```

Arguments

df	An object of class R1010 and data.frame.
breaks	An object of type string or symbol that represents columns on which to group the summarizations.
tcols	An object of type list, where each element specifies a tcol. See the examples below for details. If no tcol is provided, the count function is applied to each group on the 1010data table. A list of available functions to use with tcol can be found in the 1010data Reference Manual .
cbreaks	An object of type string or symbol that represents a list of names of the column-break columns (real or computed). All records that have the same values for these columns will be grouped together. If omitted, the result is a straight tabulation.
run	An object of type boolean that specifies whether the tabulation is to be run. If FALSE, the tabulation is added to the query in the df argument, but is not computed. The default is set to FALSE.

Details

The tabu function requires a table handle (an object of class R1010) and a set of break columns, and it returns a table handle containing the results of the tabulation.

Examples

```
## Not run:
ses <- newSession("moe1010", "P13a5e")
df <- openTable(ses, "pub.demo.retail.item")

tabu(df, breaks=c("store"), run=TRUE)

tabu(df, breaks=c(store,date), run=TRUE)

tabu(df, breaks=store, tcols=list(ucnt_sku=ucnt(sku),fd=first(date),ld=last(date)), run=TRUE)

## End(Not run)
```

tail.R1010

View the last n rows of the current R1010 table object.

Description

Allows the user to view the current state of a table (possibly transformed by a query).

Usage

```
tail.R1010(df, nrows = 20)
```

Arguments

df	An object of class R1010 and data.frame.
nrows	The number of rows from the bottom of a table to be displayed. The default is 20.

Examples

```
## Not run:
ses <- newSession("moe1010", "P13a5e")
df <- openTable(ses, "pub.demo.weather.stations")

dfq.select <- query(df,
  "<sel value=\"between(i_;5;21)\"/>")

tail(df) # shows last 20 rows of pub.demo.weather.stations
tail(dfq.select) # shows all 17 rows after <sel> operation
tail(dfq.select, 10) # shows last 10 rows after <sel> operation

## End(Not run)
```

willbe	<i>Function to create a computed column within an R1010 table.</i>
--------	--

Description

Allows the user to create a computed column within an R1010 table.

Usage

```
willbe(df, name, value, run = FALSE, ...)
```

Arguments

df	An object of class R1010 and data.frame.
name	An object of type character that will be used to name the computed column. The name may only contain alphanumeric characters or underscores and must begin with an alphabetic character (e.g., percent_total_sales). It may not contain any spaces or other special characters.
value	An object of type character that is a valid expression. See the 1010data Reference Manual .
run	An object of type boolean that specifies whether the computed column will be created. If FALSE, the willbe is added to the query in the df argument, but the computed column is not created. The default is FALSE.

Details

The willbe function requires a table handle (an object of class R1010), a name, and an expression, and it returns a table handle containing the results of the willbe (if run is set to TRUE).

Examples

```
## Not run:
ses <- newSession("larry1010", "P13a5e")
df <- openTable(ses, "pub.demo.retail.item")

willbe(df, "margin", "sales-cost", run=TRUE)

willbe(df, "store_sales", "g_sum(store; ; sales)", run=T)

willbe(df, "store_sales_by_date", "g_sum(store date; ; sales)", run=T)

## End(Not run)
```


Index

*Topic **1010data**

R1010-package, 2

clearCache, 2

dim, 3

dir (getDir), 4

getCurrMem, 3

getDir, 4

getSystemVersion, 5

head (head.R1010), 5

head.R1010, 5

info, 6

link, 7

logoutSession, 8

newSession, 8

nrow, 9

openTable, 10

posSession, 11

query, 12

R1010-package, 2

sel, 13

tabu, 14

tail (tail.R1010), 15

tail.R1010, 15

willbe, 16