Package ‘table.express’

September 7, 2019

Type  Package
Title  Build 'data.table' Expressions with Data Manipulation Verbs
Description  A specialization of 'dplyr' data manipulation verbs that parse and build expressions which are ultimately evaluated by 'data.table', letting it handle all optimizations. A set of additional verbs is also provided to facilitate some common operations on a subset of the data.
Version  0.3.1
Depends  R (>= 3.1.0)
Imports  methods, stats, utils, data.table (>= 1.9.8), dplyr, magrittr, R6, rlang (>= 0.3.1), tidyselect
Suggests  knitr, rex, rmarkdown, testthat
Date  2019-09-06
BugReports  https://github.com/asardaes/table.express/issues
License  MPL-2.0
URL  https://asardaes.github.io/table.express,
     https://github.com/asardaes/table.express
Language  en-US
Encoding  UTF-8
LazyData  TRUE
RoxygenNote  6.1.1
VignetteBuilder  knitr
Collate  'DELIMITERS-chain.R' 'DELIMITERS-end_expr.R'
        'DELIMITERS-start_expr.R' 'R6-ExprBuilder.R'
        'R6-EagerExprBuilder.R' 'UTILS-frame_append.R' 'UTILS-misc.R'
        'UTILS-nest_expr.R' 'UTILS-tidyselect.R' 'VERBS-joins.R'
        'VERBS-anti_join.R' 'VERBS-arrange.R' 'VERBS-distinct.R'
        'VERBS-filter.R' 'VERBS-filter_on.R' 'VERBS-filter_sd.R'
        'VERBS-full_join.R' 'VERBS-group_by.R' 'VERBS-inner_join.R'
        'VERBS-key_by.R' 'VERBS-left_join.R' 'VERBS-max_by.R'
        'VERBS-min_by.R' 'VERBS-mutate.R' 'VERBS-mutate_join.R'
        'VERBS-mutate_sd.R' 'VERBS-order_by.R' 'VERBS-right_join.R'
Topics documented:

'VERBS-select.R' 'VERBS-semi_join.R' 'VERBS-summarize.R'
'VERBS-transmute.R' 'VERBS-transmute_sd.R' 'VERBS-where.R'
'pkg.R'

NeedsCompilation no

Author Alexis Sarda-Espinosa [cre, aut]

Maintainer Alexis Sarda-Espinosa <alexis.sarda@gmail.com>

Repository CRAN

Date/Publication 2019-09-07 11:10:02 UTC

R topics documented:

<table>
<thead>
<tr>
<th>Function</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>table.express-package</td>
<td>3</td>
</tr>
<tr>
<td>arrange-table.express</td>
<td>5</td>
</tr>
<tr>
<td>chain</td>
<td>6</td>
</tr>
<tr>
<td>distinct-table.express</td>
<td>7</td>
</tr>
<tr>
<td>EagerExprBuilder</td>
<td>8</td>
</tr>
<tr>
<td>end_expr</td>
<td>8</td>
</tr>
<tr>
<td>ExprBuilder</td>
<td>9</td>
</tr>
<tr>
<td>extrema_by</td>
<td>10</td>
</tr>
<tr>
<td>filter-table.express</td>
<td>11</td>
</tr>
<tr>
<td>filter_on</td>
<td>12</td>
</tr>
<tr>
<td>filter_sd</td>
<td>13</td>
</tr>
<tr>
<td>frame_append</td>
<td>14</td>
</tr>
<tr>
<td>group_by-table.express</td>
<td>15</td>
</tr>
<tr>
<td>joins</td>
<td>16</td>
</tr>
<tr>
<td>key_by</td>
<td>19</td>
</tr>
<tr>
<td>mutate-table.express</td>
<td>20</td>
</tr>
<tr>
<td>mutate_sd</td>
<td>21</td>
</tr>
<tr>
<td>nest_expr</td>
<td>22</td>
</tr>
<tr>
<td>order_by-table.express</td>
<td>23</td>
</tr>
<tr>
<td>select-table.express</td>
<td>24</td>
</tr>
<tr>
<td>start_expr</td>
<td>25</td>
</tr>
<tr>
<td>summarize-table.express</td>
<td>26</td>
</tr>
<tr>
<td>transmute-table.express</td>
<td>27</td>
</tr>
<tr>
<td>transmute_sd</td>
<td>28</td>
</tr>
<tr>
<td>where-table.express</td>
<td>29</td>
</tr>
</tbody>
</table>

Index 31
Description

A specialization of `dplyr` verbs, as well as a set of custom ones, that build expressions that can be used within a `data.table`'s frame.

Note

Since this package’s functionality is based on the `rlang` package, and `rlang` is still evolving, breaking changes may be needed in the future.

Note that since version 0.3.0, it is not possible to load `table.express` and `dtplyr` at the same time, since they define the same `data.table` methods for many `dplyr` generics.

If a package uses `dplyr` without importing `data.table`, the methods in this package will try to delegate to the `data.frame` methods with a warning. To avoid the warning, use `options(table.express.warn.cedta = FALSE)`.

This software package was developed independently of any organization or institution that is or has been associated with the author.

Author(s)

Alexis Sarda-Espinosa

See Also

Useful links:

- [https://asardaes.github.io/table.express](https://asardaes.github.io/table.express)
- [https://github.com/asardaes/table.express](https://github.com/asardaes/table.express)
- Report bugs at [https://github.com/asardaes/table.express/issues](https://github.com/asardaes/table.express/issues)

Examples

```r
require("data.table")

data("mtcars")

DT <- as.data.table(mtcars)

# Simple dplyr-like transformations

DT %>%
  group_by(cyl) %>%
  filter(vs == 0, am == 1) %>%
  transmute(mean_mpg = mean(mpg)) %>%
```
```r
arrange(-cyl)

# Equivalent to previous
DT %>%
  start_expr %>%
  transmute(mean_mpg = mean(mpg)) %>%
  where(vs == 0, am == 1) %>%
  group_by(cyl) %>%
  order_by(-cyl) %>%
  end_expr

# Modification by reference
DT %>%
  where(gear % 2 != 0, carb % 2 == 0) %>%
  mutate(wt_squared = wt ^ 2)

print(DT)

# Deletion by reference
DT %>%
  mutate(wt_squared = NULL) %>%
  print

# Support for tidyslect helpers
DT %>%
  select(ends_with("t"))

# ====================================================================================
# Helpers to transform a subset of data
# Like DT[, (whole) := lapply(.SD, as.integer), .SDcols = whole]
whole <- names(DT)[sapply(DT, function(x) { all(x %% 1 == 0) })]
DT %>%
  mutate_sd(as.integer, .SDcols = whole)
sapply(DT, class)

# Like DT[, lapply(.SD, fun), .SDcols = ...]
DT %>%
  transmute_sd((.COL - mean(.COL)) / sd(.COL),
               .SDcols = setdiff(names(DT), whole))

# Filter several with the same condition
DT %>%
  filter_sd(.COL == 1, .SDcols = c("vs", "am"))

# Using secondary indices, i.e. DT[(4, 5), on = (.cyl, gear)]
DT %>%
  filter_on(cyl = 4, gear = 5) # note we don't use ==
scale_undim <- function(...) {
  as.numeric(scale(...)) # remove dimensions
}
arrange-table.express

Arrange rows

Description

Alias for order_by-table.express.

Usage

```r
## S3 method for class 'ExprBuilder'
arrange(.data, ...)
```

```r
## S3 method for class 'data.table'
arrange(.data, ...)
```
Arguments

.data An instance of ExprBuilder.

... See order_by-table.express.

Details

To see more examples, check the vignette, or the table.express-package entry.

chain

<table>
<thead>
<tr>
<th>chain</th>
<th>Chain</th>
</tr>
</thead>
</table>

Description

Build a chain of similar objects/operations.

Usage

chain(.data, ...)

## S3 method for class 'ExprBuilder'
chain(.data, ..., 
    .parent_env = rlang::caller_env())

Arguments

.data Object to be chained.

... Arguments for the specific methods.

.parent_env See end_expr().

Details

The chaining for ExprBuilder is equivalent to calling end_expr() followed by start_expr(). The ellipsis (...) is passed to both functions.

To see more examples, check the vignette, or the table.express-package entry.
Rows with distinct combinations of columns

Usage

```r
## S3 method for class 'ExprBuilder'
distinct(.data, ..., .keep = TRUE, .n = 1L, .parse = getOption("table.express.parse", FALSE))

## S3 method for class 'data.table'
distinct(.data, ...)
```

Arguments

- `.data` An instance of `ExprBuilder`.
- `...` Which columns to use to determine uniqueness.
- `.keep` See details below.
- `.n` Indices of rows to return for each unique combination of the chosen columns. See details.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.

Details

If `.keep = TRUE` (the default), the columns not mentioned in `...` are also kept. However, if a new column is created in one of the expressions therein, `.keep` can also be set to a character vector containing the names of all the columns that should be in the result in addition to the ones mentioned in `...`. See the examples.

The value of `.n` is only relevant when `.keep` is not `FALSE`. It is used to subset `.SD` in the built `data.table` expression. For example, we could get 2 rows per combination by setting `.n` to 1:2, or get the last row instead of the first by using `.N`. If more than one index is used, and not enough rows are found, some rows will have `NA`. Do note that, at least as of version 1.12.2 of `data.table`, only expressions with single indices are internally optimized.

To see more examples, check the `vignette`, or the `table.express-package` entry.

Examples

```r
data("mtcars")

# compare with .keep = TRUE
data.table::as.data.table(mtcars) %>%
```
distinct(amvs = am + vs, .keep = names(mtcars))

EagerExprBuilder

*Eager frame expression builder*

**Description**

Like `ExprBuilder`, but eager in some regards. This shouldn’t be used directly.

**Usage**

```r
EagerExprBuilder
```

**Format**

An object of class `R6ClassGenerator` of length 24.

---

end_expr

*End and evaluate expression*

**Description**

Finish the expression-building process and evaluate it.

**Usage**

```r
end_expr(.data, ...)
```

```
## S3 method for class 'ExprBuilder'
end_expr(.data, ..., .by_ref = TRUE, .parent_env)
```

**Arguments**

- `.data` The expression.
- `...` Arguments for the specific methods.
- `.by_ref` If `FALSE`, `data.table::copy()` is used before evaluation.
- `.parent_env` Optionally, the *enclosing* environment of the expression’s evaluation environment. Defaults to the caller environment.

**Details**

The `ExprBuilder` method returns a `data.table::data.table`.

To see more examples, check the `vignette`, or the `table.express-package` entry.
Frame expression builder

Description

Build an expression that will be used inside a `data.table::data.table`’s frame. This shouldn’t be used directly.

Usage

`ExprBuilder`

Format

An object of class `R6ClassGenerator` of length 24.

Fields

- `appends` Extra expressions that go at the end.
- `expr` The final expression that can be evaluated with `base::eval()` or `rlang::eval_bare()`.

Methods

- `initialize(DT, dt_pronouns = list(), .verbose)` Constructor that receives a `data.table::data.table` in `DT`. The `dt_pronouns` parameter is used internally when chaining for joins.
- `set_j(value, chain_if_needed)` Set the `j` clause expression(s), starting a new frame if the current one already has said expression set.
- `set_i(value, chain_if_needed)` Like `set_j` but for the `i` clause.
- `set_by(value, chain_if_needed)` Set the by clause expression.
- `chain(type = "frame", dt)` By default, start a new expression with the current one as its parent. If `type = "pronoun", dt` is used to start a new expression that joins the current one.
- `eval(parent_env, by_ref, ...)` Evaluate the final expression with `parent_env` as the enclosing environment. If `by_ref = FALSE`, `data.table::copy()` is called before. The ellipsis’ contents are assigned to the expression’s evaluation environment.
- `tidy_select(select_expr)` Evaluate a tidyselect call using the currently captured table.
- `print(...)` Prints the built `expr`.  

Find rows with extrema in specific columns

Description

Find rows with maxima/minima in given columns.

Usage

max_by(.data, .col, ...)

## S3 method for class 'ExprBuilder'
max_by(.data, .col, ..., .some = FALSE, .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'data.table'
max_by(.data, .col, ..., .expr = FALSE)

min_by(.data, .col, ...)

## S3 method for class 'ExprBuilder'
min_by(.data, .col, ..., .some = FALSE, .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'data.table'
min_by(.data, .col, ..., .expr = FALSE)

Arguments

.data An instance of ExprBuilder.
.col A character vector indicating the columns that will be searched for extrema.
... Optionally, columns to group by, either as characters or symbols.
.some If TRUE the rows where any of the columns specified in .col have extrema are returned.
.chain Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
.expr If the input is a data.table and .expr is TRUE, an instance of EagerExprBuilder will be returned. Useful if you want to add clauses to j, e.g. with mutate-table.express.

Details

These verbs implement the idiom shown here by leveraging nest_expr(). The whole nested expression is assigned to i in the data.table's frame. It is probably a good idea to use this on a frame that has no other frames preceding it in the current expression, given that nest_expr() uses the captured data.table, so consider using chain() when needed.
Several columns can be specified in .col, and depending on the value of .some, the rows with all or some extrema are returned, using & or | respectively. Depending on your data, using more than one column might not make sense, resulting in an empty data.table.

Examples

```r
data("mtcars")
data.table::as.data.table(mtcars) %>%
  max_by("mpg", "vs")
```

Description

Filter rows

Usage

```r
## S3 method for class 'ExprBuilder'
filter(.data, ..., .preserve)

## S3 method for class 'data.table'
filter(.data, ...)
```

Arguments

- `.data` An instance of ExprBuilder.
- `...` See where-table.express.
- `.preserve` Ignored.

Details

The ExprBuilder method is an alias for where-table.express.

The data.table::data.table method works eagerly like dplyr::filter().

To see more examples, check the vignette, or the table.express-package entry.
filter_on  
Filter with secondary indices

Description
Helper to filter specifying the on part of the data.table::data.table query.

Usage
filter_on(.data, ...)

## S3 method for class 'ExprBuilder'
filter_on(.data, ..., which = FALSE, nomatch = getOption("datatable.nomatch"), mult = "all", .negate = FALSE, .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'data.table'
filter_on(.data, ..., .expr = FALSE)

Arguments
.data An instance of ExprBuilder.
...
Key-value pairs, maybe with empty keys if the data.table already has them. See details.
which, nomatch, mult
See data.table::data.table.
.negate Whether to negate the expression and search only for rows that don’t contain the given values.
.chain Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
.expr If the input is a data.table and .expr is TRUE, an instance of EagerExprBuilder will be returned. Useful if you want to add clauses to j, e.g. with mutate-table.express.

Details
The key-value pairs in ‘...’ are processed as follows:

- The names are used as on in the data.table frame. If any name is empty, on is left missing.
- The values are packed in a list and used as i in the data.table frame.

To see more examples, check the vignette, or the table.express-package entry.
Examples

```r
data("mtcars")

data.table::as.data.table(mtcars) %>%
  filter_on(cyl = 4, gear = 5)
```

---

**filter_sd** *Filter subset of data*

Description

Helper to filter rows with the same condition applied to a subset of the data.

Usage

```r
filter_sd(.data, .SDcols, .how = Negate(is.na), ...)
```

```r
## S3 method for class 'ExprBuilder'
filter_sd(.data, .SDcols, .how = Negate(is.na),
  ..., which, .collapse = "&",
  .parse = getOption("table.express.parse", FALSE),
  .chain = getOption("table.express.chain", TRUE), .caller_env_n = 1L)
```

```r
## S3 method for class 'data.table'
filter_sd(.data, ..., .expr = FALSE)
```

Arguments

- **.data**: An instance of `ExprBuilder`.
- **.SDcols**: See `data.table::data.table` and the details here.
- **.how**: The filtering function or predicate.
- **...**: Possibly more arguments for `.how`.
- **which**: Passed to `data.table::data.table`.
- **.collapse**: See `where-table.express`.
- **.parse**: Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- **.chain**: Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
- **.caller_env_n**: Internal. Passed to `rlang::caller_env()` to find the function specified in `.how` and standardize its call.
- **.expr**: If the input is a `data.table` and `.expr` is `TRUE`, an instance of `EagerExprBuilder` will be returned. Useful if you want to add clauses to `j`, e.g. with `mutate-table.express`. 
Details

This function adds/chains an `i` expression that will be evaluated by `data.table::data.table`, and it supports the `.COL` pronoun and lambdas as formulas. The `.how` condition is applied to all `.SDcols`. Additionally, `.SDcols` supports:

- `tidyselect::select_helpers`
- A predicate using the `.COL` pronoun that should return a single logical when `.COL` is replaced by a `column` of the data.
- A formula using `.x` instead of the aforementioned `.COL`.

The caveat is that the expression is evaluated eagerly, i.e. with the currently captured `data.table`. Consider using `chain()` to explicitly capture intermediate results as actual `data.tables`.

To see more examples, check the vignette, or the `table.express-package` entry.

Examples

```r
data("mtcars")
data.table::as.data.table(mtcars) %>%
  filter_sd(c("vs", "am"), ~ .x == 1)
```

Description

Add named expressions for the `data.table::data.table` frame.

Usage

```r
frame_append(.data, ..., .parse = getOption("table.express.parse", FALSE))
```

Arguments

- `.data` An instance of `ExprBuilder`.
- `...` Expressions to add to the frame.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.

Examples

```r
data.table::data.table() %>%
  start_expr %>%
  frame_append(anything = "goes")
```
Description

Grouping by columns of a `data.table::data.table`.

Usage

```r
## S3 method for class 'ExprBuilder'
group_by(.data, ..., 
  .parse = getOption("table.express.parse", FALSE),
  .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'data.table'
group_by(.data, ...)
```

Arguments

- `.data` An instance of `ExprBuilder`.
- `...` Clause for grouping on columns. The by inside the `data.table`'s frame.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.chain` Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?

Details

Everything in `...` will be wrapped in a call to `list`.

To see more examples, check the vignette, or the `table.express-package` entry.

Examples

```r
data("mtcars")

data.table::as.data.table(mtcars) %>%
  start_expr %>%
  group_by(cyl, gear)
```
joins

Joining verbs

Description

Two-table joins. Check the "Joining verbs" vignette for more information.

Usage

```r
## S3 method for class 'ExprBuilder'
anti_join(x, y, ...)

## S3 method for class 'data.table'
anti_join(x, ..., .expr = FALSE)

## S3 method for class 'ExprBuilder'
full_join(x, y, ..., sort = TRUE, allow = TRUE, .parent_env)

## S3 method for class 'data.table'
full_join(x, ...)

## S3 method for class 'ExprBuilder'
inner_join(x, y, ...)

## S3 method for class 'data.table'
inner_join(x, ..., .expr = FALSE)

## S3 method for class 'ExprBuilder'
left_join(x, y, ..., nomatch, mult, roll, rollends, .parent_env)

## S3 method for class 'data.table'
left_join(x, y, ..., allow = FALSE, .expr = FALSE)

mutate_join(x, y, ...)

## S3 method for class 'ExprBuilder'
mutate_join(x, y, ..., .SDcols, mult, roll, rollends, allow = FALSE, .by_each = NULL, .parent_env)

## S3 method for class 'EagerExprBuilder'
mutate_join(x, ..., .parent_env = rlang::caller_env())

## S3 method for class 'data.table'
mutate_join(x, y, ...)
```
## S3 method for class 'ExprBuilder'
right_join(x, y, ..., which, nomatch, mult, roll, rollends)

## S3 method for class 'data.table'
right_join(x, ..., allow = FALSE, .expr = FALSE)

## S3 method for class 'ExprBuilder'
semi_join(x, y, ..., allow = FALSE, .eager = FALSE)

## S3 method for class 'data.table'
semi_join(x, y, ..., allow = FALSE, .eager = FALSE)

### Arguments

- **x**
  - An `ExprBuilder` instance.

- **y**
  - A `data.table::data.table` or, for some verbs (see details), a call to `nest_expr()`.

- **...**
  - Expressions for the on part of the join.

- **.expr**
  - If the input is a `data.table` and `.expr` is `TRUE`, an instance of `EagerExprBuilder` will be returned. Useful if you want to add clauses to `j`, e.g. with `mutate_table.express`.

- **sort**
  - Passed to `data.table::merge`.

- **allow**
  - Passed as `data.table`'s `allow.cartesian`.

- **.parent_env**
  - See `end_expr()`.

- **nomatch, mult, roll, rollends**
  - See `data.table::data.table`.

- **.SDcols**
  - For `mutate_join`. See the details below.

- **.by_each**
  - For `mutate_join`. See the details below.

- **which**
  - If `TRUE`, return the row numbers that matched in `x` instead of the result of the join.

- **.eager**
  - For `semi_join`. If `TRUE`, it uses `nest_expr()` to build an expression like this instead of the default one. This uses the captured `data.table` eagerly, so use `chain()` when needed. The default is lazy.

### Details

The following joins support `nest_expr()` in `y`:

- `anti_join`
- `inner_join`
- `right_join`

The `full_join` method is really a wrapper for `data.table::merge` that specifies `all = TRUE`. The expression in `x` gets evaluated, merged with `y`, and the result is captured in a new `ExprBuilder`. Useful in case you want to keep building expressions after the merge.
Mutating join

The `ExprBuilder` method for `mutate_join` implements the idiom described in this link. The columns specified in `.SDcols` are those that will be added to `x` from `y`. The specification can be done by:

- Using `tidyselect::select_helpers`.
- Passing a character vector. If the character is named, the names are taken as the new column names for the values added to `x`.
- A list, using `base::list()` or `.()`, containing:
  - Column names, either as characters or symbols.
  - Named calls expressing how the column should be summarized/modified before adding it to `x`.

The last case mentioned above is useful when the join returns many rows from `y` for each row in `x`, so they can be summarized while joining. The value of `by` in the join depends on what is passed to `.by_each`:

- If `NULL` (the default), `by` is set to `.EACHI` if a call is detected in any of the expressions from the list in `.SDcols`
- If `TRUE`, `by` is always set to `.EACHI`
- If `FALSE`, `by` is never set to `.EACHI`

See Also

data.table::data.table, dplyr::join

Examples

```r
lhs <- data.table::data.table(x = rep(c("b", "a", "c"), each = 3),
   y = c(1, 3, 6),
   v = 1:9)

rhs <- data.table::data.table(x = c("c", "b"),
   v = 8:7,
   foo = c(4, 2))

rhs %>%
   anti_join(lhs, x, v)

lhs %>%
   inner_join(rhs, x)

# creates new data.table
lhs %>%
   left_join(rhs, x)
```
# would modify lhs by reference
lhs %>%
  start_expr %>%
  mutate_join(rhs, x, .SDcols = c("foo", rhs.v = "v"))

# would modify rhs by reference, summarizing 'y' before adding it.
rhs %>%
  start_expr %>%
  mutate_join(lhs, x, .SDcols = .(y = mean(y)))

# creates new data.table
lhs %>>%
  right_join(rhs, x)

# keep only columns from lhs
lhs %>>%
  semi_join(rhs, x)

---

**key_by**

*Set key to group by*

**Description**

Group by setting key of the input.

**Usage**

```r
key_by(.data, ...)  
```

### S3 method for class 'ExprBuilder'

```r
key_by(.data, ...,  
  .parse = getOption("table.express.parse", FALSE),  
  .chain = getOption("table.express.chain", TRUE))
```

### S3 method for class 'data.table'

```r
key_by(.data, ...)
```

**Arguments**

- `.data`  
  Object to be grouped and subsequently keyed.
- `...`  
  Arguments for the specific methods.
- `.parse`  
  Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.chain`  
  Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
**Details**

Everything in ... will be wrapped in a call to list. Its contents work like Clauses for grouping on columns. The keyby inside the data.table::data.table frame.

To see more examples, check the vignette, or the table.express-package entry.

**Examples**

```r
data("mtcars")
data.table::as.data.table(mtcars) %>%
  start_expr %>%
  key_by(cyl, gear)
```

**Description**

Add or update columns of a data.table::data.table, possibly by reference using :=.

**Usage**

```r
## S3 method for class 'ExprBuilder'
mutable(data, ..., .sequential = FALSE,
         .unquote_names = TRUE, .parse = getOption("table.express.parse", FALSE), .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'EagerExprBuilder'
mutable(data, ..., .parent_env = rlang::caller_env())

## S3 method for class 'data.table'
mutable(data, ...)
```

**Arguments**

- `.data` An instance of ExprBuilder.
- `...` Mutation clauses.
- `.sequential` If TRUE, each expression in ... is assigned to a separate frame in order to enable usage of newly created columns.
- `.unquote_names` Passed to rlang::enexprs(). Set to FALSE if you want to pass the single := expression.
- `.parse` Logical. Whether to apply rlang::parse_expr() to obtain the expressions.
- `.chain` Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
- `.parent_env` See end_expr()
Details

To see more examples, check the vignette, or the table.express-package entry.

Examples

```r
data("mtcars")
data.table::as.data.table(mtcars) %>%
  start_expr %>%
  mutate(mpg_squared = mpg ^ 2)
```

---

### `mutate_sd`

**Mutate subset of data**

**Description**

Like `mutate-table.express` but possibly recycling calls.

**Usage**

```r
mutate_sd(.data, .SDcols, .how = identity, ...)
```

**Arguments**

- `.data` An instance of `ExprBuilder`.
- `.SDcols` See `data.table::data.table` and the details here.
- `.how` The function(s) or function call(s) that will perform the transformation. If many, a list should be used, either with `list()` or `.()`. If the list is named, the names will be used for the new columns’ names. Lambdas specified as formulas are supported.
- `...` Possibly more arguments for *all* functions/calls in `.how`.

---
nest_expr

<table>
<thead>
<tr>
<th>.pairwise</th>
<th>If FALSE, each function in .how is applied to each column in .SDcols (like a cartesian product).</th>
</tr>
</thead>
<tbody>
<tr>
<td>.prefix, .suffix</td>
<td>Only relevant when .how is a function: add a prefix or suffix to the new column’s name. If neither is missing, .prefix has preference.</td>
</tr>
<tr>
<td>.parse</td>
<td>Logical. Whether to apply <code>rlang::parse_expr()</code> to obtain the expressions.</td>
</tr>
<tr>
<td>.chain</td>
<td>Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?</td>
</tr>
<tr>
<td>.parent_env</td>
<td>See <code>end_expr()</code></td>
</tr>
</tbody>
</table>

**Details**

This function works similar to `transmute_sd()` but keeps all columns and _can_ modify by reference, like `mutate-table.express`. It can serve like dplyr's scoped mutation variants depending on what’s given to .SDcols.

Additionally, .SDcols supports:

- tidyselect::select_helpers
- A predicate using the .COL pronoun that should return a single logical when .COL is replaced by a _column_ of the data.
- A formula using . or .x instead of the aforementioned .COL.

The caveat is that the expression is evaluated eagerly, i.e. with the currently captured data.table. Consider using `chain()` to explicitly capture intermediate results as actual data.tables.

To see more examples, check the vignette, or the table.express-package entry.

**Examples**

```r
data("mtcars")

data.table::as.data.table(mtcars) %>%
  start_expr %>%
  mutate_sd(c("mpg", "cyl"), ~ .x * 2)
```

---

**nest_expr**

Nest expressions as a functional chain

**Description**

Nest expressions as a functional chain

**Usage**

```r
nest_expr(..., .start = TRUE, .end = .start, .parse = getOption("table.express.parse", FALSE))
```
Arguments

...  Expressions that will be part of the functional chain.
.start  Whether to add a `start_expr()` call at the beginning of the chain.
.end  Whether to add an `end_expr()` call at the end of the chain.
.parse  Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.

Details

All expressions in ... are "collapsed" with `%%`, passing the `ExprBuilder`'s captured `data.table` as the initial parameter. Names are silently dropped.

The chain is evaluated eagerly and saved in the `ExprBuilder` instance to be used during final expression evaluation.

To see more examples, check the vignette, or the `table.express-package` entry.

Description

Clause for ordering rows.

Usage

```r
order_by(.data, ...)  
```

```r
## S3 method for class 'ExprBuilder'
order_by(.data, ..., .collapse,
    .parse = getOption("table.express.parse", FALSE),
    .chain = getOption("table.express.chain", TRUE))
```

```r
## S3 method for class 'data.table'
order_by(.data, ...)
```

Arguments

.data  The input data.
...  Arguments for the specific methods.
.collapse  Ignored. See details.
.parse  Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
.chain  Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
Details

The `ExprBuilder` method dispatches to `where-table.express`, but doesn’t forward the `.collapse` argument.

To see more examples, check the vignette, or the `table.express-package` entry.

Examples

```r
data("mtcars")
data.table::as.data.table(mtcars) %>%
  order_by(-cyl, gear)
```

Description

Select columns of a `data.table::data.table`.

Usage

```r
## S3 method for class 'ExprBuilder'
select(.data, ..., .negate = FALSE,
       .parse = getOption("table.express.parse", FALSE),
       .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'EagerExprBuilder'
select(.data, ...,
       .parent_env = rlang::caller_env())

## S3 method for class 'data.table'
select(.data, ...)
```

Arguments

- `.data` An instance of `ExprBuilder`.
- `...` Clause for selecting columns. For `j` inside the `data.table`’s frame.
- `.negate` Whether to negate the selection semantics and keep only columns that do not match what’s given in `...`.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.chain` Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
- `.parent_env` See `end_expr()`
**Details**

The expressions in . . . support tidyselect::select_helpers.

To see more examples, check the vignette, or the table.express-package entry.

**Examples**

```
data("mtcars")
data.table::as.data.table(mtcars) %>%
  select(mpg:cyl)
```

---

**Description**

Start building an expression.

**Usage**

```
start_expr(.data, ...)
```

```
## S3 method for class 'data.table'
start_expr(.data, ..., .verbose = getOption("table.express.verbose", FALSE))
```

**Arguments**

- `.data` Optionally, something to capture for the expression.
- `...` Arguments for the specific methods.
- `.verbose` Whether to print more information during the expression-building process.

**Details**

The data.table::data.table method returns an ExprBuilder instance.

To see more examples, check the vignette, or the table.express-package entry.
summarize-table.express

*Summarize columns*

**Description**

Compute summaries for columns, perhaps by group.

**Usage**

```r
## S3 method for class 'ExprBuilder'
summarize(.data, ..., 
  .parse = getOption("table.express.parse", FALSE), 
  .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'EagerExprBuilder'
summarise(.data, ..., 
  .parent_env = rlang::caller_env())

## S3 method for class 'EagerExprBuilder'
summarise(.data, ..., 
  .parent_env = rlang::caller_env())

## S3 method for class 'data.table'
summarize(.data, ...) 

## S3 method for class 'data.table'
summarise(.data, ...) 
```

**Arguments**

- `.data` An instance of `ExprBuilder`.
- `...` Clauses for transmuting columns. For j inside the `data.table`'s frame.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.chain` Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
- `.parent_env` See `end_expr()`
Details

The built expression is similar to what transmute builds, but the function also checks that the results have length 1.

To see more examples, check the vignette, or the table.express-package entry.

Description

Compute and keep only new columns.

Usage

```r
## S3 method for class 'ExprBuilder'
transmute(.data, ..., .enlist = TRUE,
 .parse = getOption("table.express.parse", FALSE),
 .chain = getOption("table.express.chain", TRUE))

## S3 method for class 'EagerExprBuilder'
transmute(.data, ...,
 .parent_env = rlang::caller_env())

## S3 method for class 'data.table'
transmute(.data, ...)
```

Arguments

- `.data` An instance of `ExprBuilder`.
- `...` Clauses for transmuting columns. For `j` inside the `data.table`'s frame.
- `.enlist` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.chain` Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
- `.parent_env` See `end_expr()`

Details

Everything in `...` is wrapped in a call to `list` by default. If only one expression is given, you can set `.enlist` to `FALSE` to skip the call to `list`.

To see more examples, check the vignette, or the `table.express-package` entry.
Examples

```r
data("mtcars")

data.table::as.data.table(mtcars) %>%
  transmute(ans = mpg * 2)
```

---

transmute_sd  
Transmute subset of data

Description

Like `transmute-table.express` but for a single call and maybe specifying `.SDcols`.

Usage

```r
transmute_sd(.data, .SDcols = everything(), .how = identity, ...)
```

## S3 method for class 'ExprBuilder'
```r
transmute_sd(.data, .SDcols = everything(),
    .how = identity, ..., .parse = getOption("table.express.parse", FALSE),
    .chain = getOption("table.express.chain", TRUE))
```

## S3 method for class 'EagerExprBuilder'
```r
transmute_sd(.data, ..., 
    .parent_env = rlang::caller_env())
```

## S3 method for class 'data.table'
```r
transmute_sd(.data, ...)
```

Arguments

- `.data` An instance of `ExprBuilder`.
- `.SDcols` See `data.table::data.table` and the details here.
- `.how` The function(s) or function call(s) that will perform the transformation. If many, a list should be used, either with `list()` or `.()`. If the list is named, the names will be used for the new columns’ names. Lambdas specified as formulas are supported.
- `...` Possibly more arguments for all functions/calls in `.how`.
- `.parse` Logical. Whether to apply `rlang::parse_expr()` to obtain the expressions.
- `.chain` Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?
- `.parent_env` See `end_expr()`
Details

Like `transmute-table.express`, this function never modifies the input by reference. This function adds/chains a select expression that will be evaluated by `data.table::data.table`, possibly specifying the helper function `.transmute_matching`, which is assigned to the final expression's evaluation environment when calling `end_expr()` (i.e., `ExprBuilder`'s `eval` method).

Said function supports two pronouns that can be used by `.how` and `.SDcols:

- `.COL`: the actual values of the column.
- `.COLNAME`: the name of the column currently being evaluated.

Additionally, lambdas specified as formulas are also supported. In those cases, `.x` is equivalent to `.COL` and `.y` to `.COLNAME.

Unlike a call like `DT[, (vars) := expr]`, `.SDcols` can be created dynamically with an expression that evaluates to something that would be used in place of `vars` without eagerly using the captured `data.table`. See the examples here or in `table.express-package`.

Examples

```r
data("mtcars")

data.table::as.data.table(mtcars) %>%
  transmute_sd(~ grepl("^d", .y), ~ .x * 2)

data.table::as.data.table(mtcars) %>%
  transmute_sd(~ is.numeric(.x), ~ .x * 2)
```

where-table.express Where clause

Description

Clause for subsetting rows.

Usage

```r
where(.data, ...)
```

## S3 method for class 'ExprBuilder'

```r
where(.data, ..., which, .collapse = `&`,
  .parse = getOption("table.express.parse", FALSE),
  .chain = getOption("table.express.chain", TRUE))
```

## S3 method for class 'data.table'

```r
where(.data, ...)
```
where-table.express

Arguments

.data  The input data.
...  Arguments for the specific methods.
which  Passed to data.table::data.table.
.collapse  A boolean function which will be used to "concatenate" all conditions in ....
.parse  Logical. Whether to apply rlang::parse_expr() to obtain the expressions.
.chain  Logical. Should a new frame be automatically chained to the expression if the clause being set already exists?

Details

For ExprBuilder, the expressions in ... can call nest_expr(), and are eagerly nested if they do.
The data.table::data.table method is lazy, so it expects another verb to follow afterwards.
To see more examples, check the vignette, or the table.express-package entry.

Examples


data("mtcars")

data.table::as.data.table(mtcars) %>%
  start_expr %>%
  where(vs == 0, am == 1)


data.table::as.data.table(mtcars) %>%
  where(vs == 0) %>%
  transmute(mpg = round(mpg))
Index

*Topic* datasets
  EagerExprBuilder, 8
  ExprBuilder, 9
  :=, 20
  %>%, 23
  anti_join.data.table (joins), 16
  anti_join.ExprBuilder (joins), 16
  arrange.data.table
    (arrange-table.express), 5
  arrange.ExprBuilder
    (arrange-table.express), 5
  base::eval(), 9
  base::list(), 18
  chain, 6
  chain(), 10, 14, 17, 22
  data.table, 3, 17
  data.table::copy(), 8, 9
  data.table::data.table, 8, 9, 11–15, 17, 18, 20, 21, 24, 25, 28–30
  data.table::merge, 17
  distinct-table.express, 7
  distinct.data.table
    (distinct-table.express), 7
  distinct.ExprBuilder
    (distinct-table.express), 7
  dplyr, 3
  dplyr’s scoped mutation variants, 22
  dplyr::filter(), 11
  dplyr::join, 18
  EagerExprBuilder, 8, 10, 12, 13, 17
  end_expr, 8
  end_expr(), 6, 17, 20, 22–24, 26–29
  ExprBuilder, 6–8, 9, 10–15, 17, 18, 20, 21, 23–30
  extrema_by, 10
  filter-table.express, 11
  filter.data.table
    (filter-table.express), 11
  filter.ExprBuilder
    (filter-table.express), 11
  filter_on, 12
  filter_sd, 13
  frame_append, 14
  full_join.data.table (joins), 16
  full_join.ExprBuilder (joins), 16
  group_by-table.express, 15
  group_by.data.table
    (group_by-table.express), 15
  group_by.ExprBuilder
    (group_by-table.express), 15
  inner_join.data.table (joins), 16
  inner_join.ExprBuilder (joins), 16
  joins, 16
  key_by, 19
  left_join.data.table (joins), 16
  left_join.ExprBuilder (joins), 16
  max_by (extrema_by), 10
  min_by (extrema_by), 10
  mutate-table.express, 10, 12, 13, 17, 20, 21, 22
  mutate.data.table
    (mutate-table.express), 20
  mutate.ExprBuilder
    (mutate-table.express), 20
  mutate.EagerExprBuilder
    (mutate-table.express), 20
  mutate_join (joins), 16
  mutate_sd, 21
  nest_expr, 22
nest_expr(), 10, 17, 30

order_by (order_by-table.express), 23
order_by-table.express, 5, 6, 23
order_by-data.table
  (order_by-table.express), 23
order_by-ExprBuilder
  (order_by-table.express), 23
right_join-data.table (joins), 16
right_join-ExprBuilder (joins), 16
rlang::caller_env(), 13
rlang::enexprs(), 20
rlang::eval_bare(), 9
rlang::parse_expr(), 7, 13–15, 19, 20, 22–24, 26–28, 30
select-table.express, 24
select-data.table
  (select-table.express), 24
select-EagerExprBuilder
  (select-table.express), 24
select-ExprBuilder
  (select-table.express), 24
semi_join-data.table (joins), 16
semi_join-ExprBuilder (joins), 16
standardize, 13
start_expr, 25
start_expr(), 6, 23
summarise-data.table
  (summarize-table.express), 26
summarise-EagerExprBuilder
  (summarize-table.express), 26
summarise-ExprBuilder
  (summarize-table.express), 26
summarize-table.express, 26
summarize-data.table
  (summarize-table.express), 26
summarize-EagerExprBuilder
  (summarize-table.express), 26
summarize-ExprBuilder
  (summarize-table.express), 26
table.express (table.express-package), 3
table.express-package, 3, 6–8, 11, 12, 14, 15, 20–25, 27, 29, 30
tidyselect::select_helpers, 14, 18, 22, 25
transmute-table.express, 27, 28, 29