

Package ‘drake’

August 19, 2019

Title A Pipeline Toolkit for Reproducible Computation at Scale

Version 7.6.1

Description A general-purpose computational engine for data analysis, drake rebuilds intermediate data objects when their dependencies change, and it skips work when the results are already up to date. Not every execution starts from scratch, there is native support for parallel and distributed computing, and completed projects have tangible evidence that they are reproducible. Extensive documentation, from beginner-friendly tutorials to practical examples and more, is available at the reference website [<https://ropensci.github.io/drake/>](https://ropensci.github.io/drake/) and the online manual [<https://ropenscilabs.github.io/drake-manual/>](https://ropenscilabs.github.io/drake-manual/).

License GPL-3

URL <https://github.com/ropensci/drake>

BugReports <https://github.com/ropensci/drake/issues>

Depends R (>= 3.3.0)

Imports base64url,
digest,
igraph,
methods,
rlang (>= 0.2.0),
storr (>= 1.1.0),
txtq (>= 0.1.3),
utils

Suggests abind,
bindr,
callr,
cli (>= 1.1.0),
clustermq (>= 0.8.8),
crayon,
curl (>= 2.7),
datasets,
downloader,
fst,
future,
ggplot2,
ggraph,

grDevices,
 keras,
 knitr,
 lubridate,
 networkD3,
 parallel,
 prettycode,
 Rcpp,
 rmarkdown,
 rstudioapi,
 stats,
 styler,
 testthat ($\geq 2.1.0$),
 tibble,
 tidyselect ($\geq 0.2.4$),
 txtplot,
 usethis,
 visNetwork,
 webshot

VignetteBuilder knitr

Encoding UTF-8

Language en-US

Roxygen list(markdown = TRUE)

RoxygenNote 6.1.1

R topics documented:

drake-package	4
bind_plans	4
build_times	5
cached	6
clean	8
clean_mtcars_example	9
code_to_plan	10
combine	11
cross	13
deps_code	15
deps_knitr	16
deps_profile	16
deps_target	17
diagnose	18
drake_build	19
drake_cache	21
drake_cache_log	22
drake_config	24
drake_debug	31
drake_envir	32
drake_example	34
drake_examples	35
drake_gc	35

drake_get_session_info	37
drake_ggraph	38
drake_graph_info	39
drake_history	41
drake_hpc_template_file	43
drake_hpc_template_files	44
drake_plan	45
drake_plan_source	49
drake_slice	50
expose_imports	51
failed	53
file_in	54
file_out	56
file_store	58
find_cache	59
ignore	59
knitr_in	61
legend_nodes	63
load_mtcars_example	64
make	65
map	73
missed	75
new_cache	76
no_deps	77
outdated	78
plan_to_code	79
plan_to_notebook	80
predict_runtime	81
predict_workers	83
progress	84
readd	86
read_drake_seed	88
recoverable	90
render_drake_ggraph	91
render_drake_graph	92
render_sankey_drake_graph	94
render_text_drake_graph	95
rescue_cache	96
running	97
r_make	98
sankey_drake_graph	101
show_source	103
split	103
target	105
text_drake_graph	107
tracked	109
transform_plan	110
trigger	112
use_drake	113
vis_drake_graph	114

drake-package

drake: A pipeline toolkit for reproducible computation at scale.

Description

drake is a pipeline toolkit (<https://github.com/pditommaso/awesome-pipeline>) and a scalable, R-focused solution for reproducibility and high-performance computing.

Author(s)

William Michael Landau <will.landau@gmail.com>

References

<https://github.com/ropensci/drake>

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    library(drake)
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Build everything.
    make(my_plan) # Nothing is done because everything is already up to date.
    reg2 = function(d) { # Change one of your functions.
      d$x3 = d$x^3
      lm(y ~ x3, data = d)
    }
    make(my_plan) # Only the pieces depending on reg2() get rebuilt.
    # Write a flat text log file this time.
    make(my_plan, cache_log_file = TRUE)
    # Read/load from the cache.
    readd(small)
    loadadd(large)
    head(large)
  }
})

## End(Not run)
```

bind_plans

Row-bind together drake plans **Stable**

Description

Combine drake plans together in a way that correctly fills in missing entries.

Usage

```
bind_plans(...)
```

Arguments

... Workflow plan data frames (see [drake_plan\(\)](#)).

See Also

[drake_plan\(\)](#), [make\(\)](#)

Examples

```
# You might need to refresh your data regularly (see ?triggers).
download_plan <- drake_plan(
  data = target(
    command = download_data(),
    trigger = "always"
  )
)
# But if the data don't change, the analyses don't need to change.
analysis_plan <- drake_plan(
  usage = get_usage_metrics(data),
  topline = scrape_toplevel_table(data)
)
your_plan <- bind_plans(download_plan, analysis_plan)
your_plan
```

build_times

*See the time it took to build each target. **Maturing***

Description

Applies to targets in your plan, not imports or files.

Usage

```
build_times(..., path = NULL, search = NULL, digits = 3,
  cache = drake_cache(path = path, verbose = verbose),
  targets_only = NULL, verbose = 1L, jobs = 1, type = c("build",
    "command"), list = character(0))
```

Arguments

...	Targets to load from the cache: as names (symbols) or character strings. If the <code>tidyselect</code> package is installed, you can also supply <code>dplyr</code> -style <code>tidyselect</code> commands such as <code>starts_with()</code> , <code>ends_with()</code> , and <code>one_of()</code> .
path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
search	Deprecated.
digits	How many digits to round the times to.
cache	drake cache. See new_cache() . If supplied, <code>path</code> is ignored.
targets_only	Deprecated.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing.

	<ul style="list-style-type: none"> • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
jobs	Number of jobs/workers for parallel processing.
type	Type of time you want: either "build" for the full build time including the time it took to store the target, or "command" for the time it took just to run the command.
list	Character vector of targets to select.

Value

A data frame of times, each from `system.time()`.

See Also

`predict_runtime()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr")) {
    if (requireNamespace("lubridate")) {
      # Show the build times for the mtcars example.
      load_mtcars_example() # Get the code with drake_example("mtcars").
      make(my_plan) # Build all the targets.
      print(build_times()) # Show how long it took to build each target.
    }
  }
})

## End(Not run)
```

cached

List targets in the cache. **Maturing**

Description

Tip: read/load a cached item with `readd()` or `loadd()`.

Usage

```
cached(..., list = character(0), no_imported_objects = FALSE,
  path = NULL, search = NULL, cache = drake_cache(path = path,
  verbose = verbose), verbose = 1L, namespace = NULL, jobs = 1,
  targets_only = TRUE)
```

Arguments

<code>...</code>	Deprecated. Do not use. Objects to load from the cache, as names (unquoted) or character strings (quoted). Similar to <code>...</code> in <code>remove()</code> .
<code>list</code>	Deprecated. Do not use. Character vector naming objects to be loaded from the cache. Similar to the <code>list</code> argument of <code>remove()</code> .
<code>no_imported_objects</code>	Logical, deprecated. Use <code>targets_only</code> instead.
<code>path</code>	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
<code>search</code>	Deprecated.
<code>cache</code>	drake cache. See <code>new_cache()</code> . If supplied, <code>path</code> is ignored.
<code>verbose</code>	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
<code>namespace</code>	Character scalar, name of the storrr namespace to use for listing objects.
<code>jobs</code>	Number of jobs/workers for parallel processing.
<code>targets_only</code>	Logical. If <code>TRUE</code> just list the targets. If <code>FALSE</code> , list files and imported objects too.

Value

Either a named logical indicating whether the given targets or cached or a character vector listing all cached items, depending on whether any targets are specified.

See Also

`readd()`, `loadadd()`, `drake_plan()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    if (requireNamespace("lubridate")) {
      load_mtcars_example() # Load drake's canonical example.
      make(my_plan) # Run the project, build all the targets.
      cached()
      cached(targets_only = FALSE)
    }
  }
})

## End(Not run)
```

clean	<i>Invalidate and deregister targets. Maturing</i>
-------	---

Description

Force targets to be out of date and remove target names from the data in the cache.

Usage

```
clean(..., list = character(0), destroy = FALSE, path = NULL,
  search = NULL, cache = drake::drake_cache(path = path, verbose =
  verbose), verbose = 1L, jobs = 1, force = FALSE,
  garbage_collection = FALSE, purge = FALSE)
```

Arguments

...	Targets to remove from the cache: as names (symbols) or character strings. If the <code>tidyselect</code> package is installed, you can also supply <code>dplyr</code> -style <code>tidyselect</code> commands such as <code>starts_with()</code> , <code>ends_with()</code> , and <code>one_of()</code> .
list	Character vector naming targets to be removed from the cache. Similar to the <code>list</code> argument of <code>remove()</code> .
destroy	Logical, whether to totally remove the drake cache. If <code>destroy</code> is <code>FALSE</code> , only the targets from <code>make()</code> are removed. If <code>TRUE</code> , the whole cache is removed, including session metadata, etc.
path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
search	Deprecated.
cache	drake cache. See <code>new_cache()</code> . If supplied, <code>path</code> is ignored.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
jobs	Number of jobs for light parallelism (disabled on Windows).
force	Logical, whether to try to clean the cache even though the project may not be back compatible with the current version of drake.
garbage_collection	Logical, whether to call <code>cache\$gc()</code> to do garbage collection. If <code>TRUE</code> , cached data with no remaining references will be removed. This will slow down <code>clean()</code> , but the cache could take up far less space afterwards. See the <code>gc()</code> method for <code>storr</code> caches.
purge	Logical, whether to remove objects from metadata namespaces such as "meta", "build_times", and "errors".

Details

By default, `clean()` invalidates **all** targets, so be careful. `clean()` always:

1. Forces targets to be out of date so the next `make()` does not skip them.
2. Deregisters targets so `loadadd(your_target)` and `readadd(your_target)` no longer work.

By default, `clean()` does not actually remove the underlying data. Even old targets from the distant past are still in the cache and recoverable via `drake_history()` and `make(recover = TRUE)`. To actually remove target data from the cache, as well as any `file_out()` files from any targets you are currently cleaning, run `clean(garbage_collection = TRUE)`. Garbage collection is slow, but it reduces the storage burden of the cache.

Value

Invisibly return `NULL`.

See Also

`drake_gc()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    # Show all registered targets in the cache.
    cached()
    # Deregister 'summ_regression1_large' and 'small' in the cache.
    clean(summ_regression1_large, small)
    # Those objects are no longer registered as targets.
    cached()
    # Rebuild the invalidated/outdated targets.
    make(my_plan)
    # Clean everything.
    clean()
    # But the data objects and files are not actually gone!
    file.exists("report.md")
    drake_history()
    make(my_plan, recover = TRUE)
    # You need garbage collection to actually remove the data
    # and any file_out() files of any uncleaned targets.
    clean(garbage_collection = TRUE)
    drake_history()
    make(my_plan, recover = TRUE)
  }
})

## End(Not run)
```

`clean_mtcars_example` *Clean the mtcars example from `drake_example("mtcars")`* **Stable**

Description

This function deletes files. Use at your own risk. Destroys the `.drake/` cache and the `report.Rmd` file in the current working directory. Your working directory (`getcwd()`) must be the folder from which you first ran `load_mtcars_example()` and `make(my_plan)`.

Usage

```
clean_mtcars_example()
```

Value

nothing

See Also

[load_mtcars_example\(\)](#), [clean\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    # Populate your workspace and write 'report.Rmd'.
    load_mtcars_example() # Get the code: drake_example("mtcars")
    # Check the dependencies of an imported function.
    deps_code(reg1)
    # Check the dependencies of commands in the workflow plan.
    deps_code(my_plan$command[1])
    deps_code(my_plan$command[4])
    # Plot the interactive network visualization of the workflow.
    config <- drake_config(my_plan)
    outdated(config) # Which targets are out of date?
    # Run the workflow to build all the targets in the plan.
    make(my_plan)
    outdated(config) # Everything should be up to date.
    # For the reg2() model on the small dataset,
    # the p-value is so small that there may be an association
    # between weight and fuel efficiency after all.
    readd(coef_regression2_small)
    # Clean up the example.
    clean_mtcars_example()
  }
})

## End(Not run)
```

code_to_plan

Turn an R script file or knitr / R Markdown report into a drake plan.
Stable

Description

[code_to_plan\(\)](#), [plan_to_code\(\)](#), and [plan_to_notebook\(\)](#) together illustrate the relationships between drake plans, R scripts, and R Markdown documents.

Usage

```
code_to_plan(path)
```

Arguments

path A file path to an R script or knitr report.

Details

This feature is easy to break, so there are some rules for your code file:

1. Stick to assigning a single expression to a single target at a time. For multi-line commands, please enclose the whole command in curly braces. Conversely, compound assignment is not supported (e.g. `target_1 <-target_2 <-target_3 <-get_data()`).
2. Once you assign an expression to a variable, do not modify the variable any more. The target/command binding should be permanent.
3. Keep it simple. Please use the assignment operators rather than `assign()` and similar functions.

See Also

[drake_plan\(\)](#), [make\(\)](#), [plan_to_code\(\)](#), [plan_to_notebook\(\)](#)

Examples

```
plan <- drake_plan(
  raw_data = read_excel(file_in("raw_data.xlsx")),
  data = raw_data,
  hist = create_plot(data),
  fit = lm(Sepal.Width ~ Petal.Width + Species, data)
)
file <- tempfile()
# Turn the plan into an R script at the given file path.
plan_to_code(plan, file)
# Here is what the script looks like.
cat(readLines(file), sep = "\n")
# Convert back to a drake plan.
code_to_plan(file)
```

combine

Define aggregates of other targets **Maturing**

Description

Similar `summarize()`, from `dplyr`, except it defines new targets in drake.

Usage

```
combine(..., .by, .id, .tag_in, .tag_out)
```

Arguments

...	Grouping variables. New grouping variables must be supplied with their names and values, existing grouping variables can be given as symbols without any values assigned.
.by	Symbol or vector of symbols of grouping variables. <code>combine()</code> aggregates/groups targets by the grouping variables in <code>.by</code>
.id	Symbol or vector of symbols naming grouping variables to incorporate into target names. Useful for creating short target names. Set <code>.id = FALSE</code> to use integer indices as target name suffixes.
.tag_in	A symbol or vector of symbols. Tags assign targets to grouping variables. Use <code>.tag_in</code> to assign <i>untransformed</i> targets to grouping variables.
.tag_out	Just like <code>.tag_in</code> , except that <code>.tag_out</code> assigns <i>transformed</i> targets to grouping variables.

Details

Only valid within a call to `target()` in `drake_plan()`. See the examples below.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`map`, `split`, `cross`, `drake_plan`, `target`

Examples

```
models <- c("glm", "hierarchical")
plan <- drake_plan(
  data = target(
    get_data(x),
    transform = map(x = c("simulated", "survey"))
  ),
  analysis = target(
    analyze_data(data, model),
    transform = cross(data, model = !!models, .id = c(x, model))
  ),
  summary = target(
    summarize_analysis(analysis),
    transform = map(analysis, .id = c(x, model))
  ),
  results = target(
    bind_rows(summary),
    transform = combine(summary, .by = data)
  )
)
plan
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}
```

cross

Define a target for each combination of values **Maturing**

Description

Similar `crossing()`, from `tidyr`, except it defines new targets in `drake`.

Usage

```
cross(..., .data, .id, .tag_in, .tag_out)
```

Arguments

<code>...</code>	Grouping variables. New grouping variables must be supplied with their names and values, existing grouping variables can be given as symbols without any values assigned.
<code>.data</code>	A data frame of new grouping variables with grouping variable names as column names and values as elements.
<code>.id</code>	Symbol or vector of symbols naming grouping variables to incorporate into target names. Useful for creating short target names. Set <code>.id = FALSE</code> to use integer indices as target name suffixes.
<code>.tag_in</code>	A symbol or vector of symbols. Tags assign targets to grouping variables. Use <code>.tag_in</code> to assign <i>untransformed</i> targets to grouping variables.
<code>.tag_out</code>	Just like <code>.tag_in</code> , except that <code>.tag_out</code> assigns <i>transformed</i> targets to grouping variables.

Details

Only valid within a call to `target()` in `drake_plan()`. See the examples below.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`map`, `split`, `combine`, `drake_plan`, `target`

Examples

```
models <- c("glm", "hierarchical")
plan <- drake_plan(
  data = target(
    get_data(x),
    transform = map(x = c("simulated", "survey"))
  ),
  analysis = target(
    analyze_data(data, model),
    transform = cross(data, model = !!models, .id = c(x, model))
  ),
  summary = target(
    summarize_analysis(analysis),
    transform = map(analysis, .id = c(x, model))
  ),
  results = target(
    bind_rows(summary),
    transform = combine(summary, .by = data)
  )
)
```

```

)
plan
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}

```

deps_code

List the dependencies of a function or command **Stable****Description**

Functions are assumed to be imported, and language/text are assumed to be commands in a plan.

Usage

```
deps_code(x)
```

Arguments

x A function, expression, or text.

Value

A data frame of the dependencies.

See Also

[deps_target\(\)](#), [deps_knitr\(\)](#)

Examples

```

# Your workflow likely depends on functions in your workspace.
f <- function(x, y) {
  out <- x + y + g(x)
  saveRDS(out, "out.rds")
}
# Find the dependencies of f. These could be R objects/functions
# in your workspace or packages. Any file names or target names
# will be ignored.
deps_code(f)
# Define a workflow plan data frame that uses your function f().
my_plan <- drake_plan(
  x = 1 + some_object,
  my_target = x + readRDS(file_in("tracked_input_file.rds")),
  return_value = f(x, y, g(z + w))
)
# Get the dependencies of workflow plan commands.
# Here, the dependencies could be R functions/objects from your workspace
# or packages, imported files, or other targets in the workflow plan.
deps_code(my_plan$command[[1]])
deps_code(my_plan$command[[2]])
deps_code(my_plan$command[[3]])
# You can also supply expressions or text.
deps_code(quote(x + y + 123))
deps_code("x + y + 123")

```

deps_knitr	<i>Find the drake dependencies of a dynamic knitr report target. Stable</i>
------------	--

Description

Dependencies in knitr reports are marked by `loadadd()` and `readadd()` in active code chunks.

Usage

```
deps_knitr(path)
```

Arguments

path	Encoded file path to the knitr/R Markdown document. Wrap paths in <code>file_store()</code> to encode.
------	--

Value

A data frame of dependencies.

See Also

`deps_code()`, `deps_target()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  load_mtcars_example() # Get the code with drake_example("mtcars").
  deps_knitr("report.Rmd")
})

## End(Not run)
```

deps_profile	<i>Find out why a target is out of date. Stable</i>
--------------	--

Description

The dependency profile can give you a hint as to why a target is out of date. It can tell you if

- the command changed (`deps_profile()` reports the *hash* of the command, not the command itself)
- at least one input file changed,
- at least one output file changed,
- or a non-file dependency changed. For this last part, the imports need to be up to date in the cache, which you can do with `outdated()` or `make(skip_targets = TRUE)`.
- the pseudo-random number generator seed changed. Unfortunately, `deps_profile()` does not currently get more specific than that.

Usage

```
deps_profile(target, config, character_only = FALSE)
```

Arguments

`target` Name of the target.

`config` Configuration list output by `drake_config()` or `make()`.

`character_only` Logical, whether to assume `target` is a character string rather than a symbol.

Value

A data frame of old and new values for each of the main triggers, along with an indication of which values changed since the last `make()`.

See Also

`diagnose()`, `deps_code()`, `make()`, `drake_config()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Load drake's canonical example.
    make(my_plan) # Run the project, build the targets.
    config <- drake_config(my_plan)
    # Get some example dependency profiles of targets.
    deps_profile(small, config = config)
    # Change a dependency.
    simulate <- function(x) {}
    # Update the in-memory imports in the cache
    # so deps_profile can detect changes to them.
    # Changes to targets are already cached.
    make(my_plan, skip_targets = TRUE)
    # The dependency hash changed.
    deps_profile(small, config = config)
  }
})

## End(Not run)
```

deps_target	<i>List the dependencies of a target</i> Stable
-------------	--

Description

Intended for debugging and checking your project. The dependency structure of the components of your analysis decides which targets are built and when.

Usage

```
deps_target(target, config, character_only = FALSE)
```

Arguments

target	A symbol denoting a target name, or if character_only is TRUE, a character scalar denoting a target name.
config	An output list from <code>drake_config()</code> .
character_only	Logical, whether to assume target is a character string rather than a symbol.

Value

A data frame with the dependencies listed by type (globals, files, etc).

See Also

`deps_code()`, `deps_knitr()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  load_mtcars_example() # Get the code with drake_example("mtcars").
  config <- drake_config(my_plan)
  deps_target("regression1_small", config = config)
})

## End(Not run)
```

diagnose

Get diagnostic metadata on a target. **Maturing**

Description

Diagnostics include errors, warnings, messages, runtimes, and other context/metadata from when a target was built or an import was processed. If your target's last build succeeded, then `diagnose(your_target)` has the most current information from that build. But if your target failed, then only `diagnose(your_target)$error`, `diagnose(your_target)$warnings`, and `diagnose(your_target)$messages` correspond to the failure, and all the other metadata correspond to the last build that completed without an error.

Usage

```
diagnose(target = NULL, character_only = FALSE, path = NULL,
  search = NULL, cache = drake::drake_cache(path = path, verbose =
  verbose), verbose = 1L)
```

Arguments

target	Name of the target of the error to get. Can be a symbol if character_only is FALSE, must be a character if character_only is TRUE.
character_only	Logical, whether target should be treated as a character or a symbol. Just like <code>character.only</code> in <code>library()</code> .
path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or NULL.
search	Deprecated.

cache	drake cache. See new_cache() . If supplied, path is ignored.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.

Value

Either a character vector of target names or an object of class "error".

See Also

[failed\(\)](#), [progress\(\)](#), [readd\(\)](#), [drake_plan\(\)](#), [make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  diagnose() # List all the targets with recorded error logs.
  # Define a function doomed to failure.
  f <- function() {
    stop("unusual error")
  }
  # Create a workflow plan doomed to failure.
  bad_plan <- drake_plan(my_target = f())
  # Running the project should generate an error
  # when trying to build 'my_target'.
  try(make(bad_plan), silent = FALSE)
  failed() # List the failed targets from the last make() (my_target).
  # List targets that failed at one point or another
  # over the course of the project (my_target).
  # drake keeps all the error logs.
  diagnose()
  # Get the error log, an object of class "error".
  error <- diagnose(my_target)$error # See also warnings and messages.
  str(error) # See what's inside the error log.
  error$calls # View the traceback. (See the traceback() function).
})

## End(Not run)
```

drake_build

*Build/process a single target or import. **Maturing***

Description

Also load the target's dependencies beforehand.

Usage

```
drake_build(target, config = NULL, meta = NULL,
  character_only = FALSE, envir = NULL, jobs = 1, replace = FALSE)
```

Arguments

target	Name of the target.
config	Internal configuration list.
meta	Deprecated.
character_only	Logical, whether name should be treated as a character or a symbol (just like <code>character.only</code> in <code>library()</code>).
envir	Environment to load objects into. Defaults to the calling environment (current workspace).
jobs	Number of parallel jobs for loading objects. On non-Windows systems, the loading process for multiple objects can be lightly parallelized via <code>parallel::mclapply()</code> . just set jobs to be an integer greater than 1. On Windows, jobs is automatically demoted to 1.
replace	Logical. If FALSE, items already in your environment will not be replaced.

Value

The value of the target right after it is built.

See Also

[drake_debug\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    # This example is not really a user-side demonstration.
    # It just walks through a dive into the internals.
    # Populate your workspace and write 'report.Rmd'.
    load_mtcars_example() # Get the code with drake_example("mtcars").
    # Create the master internal configuration list.
    config <- drake_config(my_plan)
    out <- drake_build(small, config = config)
    # Now includes `small`.
    cached()
    head(readd(small))
    # `small` was invisibly returned.
    head(out)
    # If you previously called make(),
    # `config` is just read from the cache.
    make(my_plan, verbose = FALSE)
    config <- drake_config(my_plan)
    result <- drake_build(small, config = config)
    head(result)
  }
})

## End(Not run)
```

drake_cache

Get the cache of a drake project. **Stable**

Description

`make()` saves the values of your targets so you rarely need to think about output files. By default, the cache is a hidden folder called `.drake/`. You can also supply your own storrr cache to the `cache` argument of `make()`. The `drake_cache()` function retrieves this cache.

Usage

```
drake_cache(path = NULL, verbose = 1L, console_log_file = NULL)
```

Arguments

- | | |
|------------------|---|
| path | Character. Set path to the path of a <code>storr::storr_rds()</code> cache to retrieve a specific cache generated by <code>storr::storr_rds()</code> or <code>drake::new_cache()</code> . If the path argument is NULL, <code>drake_cache()</code> searches up through parent directories to find a folder called <code>.drake/</code> . |
| verbose | Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway. |
| console_log_file | Optional character scalar of a file name or connection object (such as <code>stdout()</code>) to dump maximally verbose log information for <code>make()</code> and other functions (all functions that accept a <code>config</code> argument, plus <code>drake_config()</code>). If you choose to use a text file as the console log, it will persist over multiple function calls until you delete it manually. Fields in each row the log file, from left to right: - The node name (short host name) of the computer (from <code>Sys.info()["nodename"]</code>). - The process ID (from <code>Sys.getpid()</code>). - A timestamp with the date and time (in microseconds). - A brief description of what drake was doing. The fields are separated by pipe symbols (<code>" "</code>). |

Details

`drake_cache()` actually returns a *decorated* storrr, an object that *contains* a storrr (plus bells and whistles). To get the *actual* inner storrr, use `drake_cache()$storr`.

Value

A drake/storrr cache in a folder called `.drake/`, if available. NULL otherwise.

See Also

`new_cache()`, `drake_config()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    clean(destroy = TRUE)
    # No cache is available.
    drake_cache() # NULL
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    x <- drake_cache() # Now, there is a cache.
    y <- storr::storr_rds(".drake") # Nearly equivalent.
    # List the objects readable from the cache with readd().
    x$list()
    # drake_cache() actually returns a *decorated* storr.
    # The *real* storr is inside.
    drake_cache()$storr
  }
})

## End(Not run)
```

drake_cache_log

*Get the state of the cache. **Maturing***

Description

Get the fingerprints of all the targets in a data frame. This functionality is like `make(..., cache_log_file = TRUE)`, but separated and more customizable. Hopefully, this functionality is a step toward better data versioning tools.

Usage

```
drake_cache_log(path = NULL, search = NULL,
  cache = drake::drake_cache(path = path, verbose = verbose),
  verbose = 1L, jobs = 1, targets_only = FALSE)
```

Arguments

path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
search	Deprecated.
cache	drake cache. See new_cache() . If supplied, path is ignored.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
jobs	Number of jobs/workers for parallel processing.
targets_only	Logical, whether to output information only on the targets in your workflow plan data frame. If <code>targets_only</code> is <code>FALSE</code> , the output will include the hashes of both targets and imports.

Details

A hash is a fingerprint of an object's value. Together, the hash keys of all your targets and imports represent the state of your project. Use `drake_cache_log()` to generate a data frame with the hash keys of all the targets and imports stored in your cache. This function is particularly useful if you are storing your drake project in a version control repository. The cache has a lot of tiny files, so you should not put it under version control. Instead, save the output of `drake_cache_log()` as a text file after each `make()`, and put the text file under version control. That way, you have a changelog of your project's results. See the examples below for details. Depending on your project's history, the targets may be different than the ones in your workflow plan data frame. Also, the keys depend on the hash algorithm of your cache. To define your own hash algorithm, you can create your own storrr cache and give it a hash algorithm (e.g. `storrr_rds(hash_algorithm = "murmur32")`)

Value

Data frame of the hash keys of the targets and imports in the cache

See Also

`cached()`, `drake_cache()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    # Load drake's canonical example.
    load_mtcars_example() # Get the code with drake_example()
    # Run the project, build all the targets.
    make(my_plan)
    # Get a data frame of all the hash keys.
    # If you want a changelog, be sure to do this after every make().
    cache_log <- drake_cache_log()
    head(cache_log)
    # Suppress partial arg match warnings.
    suppressWarnings(
      # Save the hash log as a flat text file.
      write.table(
        x = cache_log,
        file = "drake_cache.log",
        quote = FALSE,
        row.names = FALSE
      )
    )
    # At this point, put drake_cache.log under version control
    # (e.g. with 'git add drake_cache.log') alongside your code.
    # Now, every time you run your project, your commit history
    # of hash_log.txt is a changelog of the project's results.
    # It shows which targets and imports changed on every commit.
    # It is extremely difficult to track your results this way
    # by putting the raw '.drake/' cache itself under version control.
  }
})

## End(Not run)
```

Description

`drake_config()` does all the preprocessing that `make()` needs to build targets. This includes interpreting the plan (from `drake_plan()`) and analyzing how all the targets fit together. The result is a list of objects that `make()` needs to keep track of everything during runtime.

Usage

```
drake_config(plan, targets = NULL, envir = parent.frame(),
  verbose = 1L, hook = NULL, cache = drake::drake_cache(verbose =
  verbose, console_log_file = console_log_file), fetch_cache = NULL,
  parallelism = "loop", jobs = 1L, jobs_preprocess = 1L,
  packages = rev(.packages()), lib_loc = NULL,
  prework = character(0), prepend = NULL, command = NULL,
  args = NULL, recipe_command = NULL, timeout = NULL, cpu = Inf,
  elapsed = Inf, retries = 0, force = FALSE, log_progress = TRUE,
  graph = NULL, trigger = drake::trigger(), skip_targets = FALSE,
  skip_imports = FALSE, skip_safety_checks = FALSE,
  lazy_load = "eager", session_info = TRUE, cache_log_file = NULL,
  seed = NULL, caching = c("master", "worker"), keep_going = FALSE,
  session = NULL, pruning_strategy = NULL, makefile_path = NULL,
  console_log_file = NULL, ensure_workers = NULL,
  garbage_collection = FALSE, template = list(), sleep = function(i)
  0.01, hasty_build = NULL, memory_strategy = "speed", layout = NULL,
  lock_envir = TRUE, history = TRUE, recover = FALSE,
  recoverable = TRUE, curl_handles = list())
```

Arguments

<code>plan</code>	Workflow plan data frame. A workflow plan data frame is a data frame with a target column and a command column. (See the details in the <code>drake_plan()</code> help file for descriptions of the optional columns.) Targets are the objects that drake generates, and commands are the pieces of R code that produce them. You can create and track custom files along the way (see <code>file_in()</code> , <code>file_out()</code> , and <code>knitr_in()</code>). Use the function <code>drake_plan()</code> to generate workflow plan data frames.
<code>targets</code>	Character vector, names of targets to build. Dependencies are built too. Together, the plan and targets comprise the workflow network (i.e. the graph argument). Changing either will change the network.
<code>envir</code>	Environment to use. Defaults to the current workspace, so you should not need to worry about this most of the time. A deep copy of <code>envir</code> is made, so you don't need to worry about your workspace being modified by <code>make</code> . The deep copy inherits from the global environment. Wherever necessary, objects and functions are imported from <code>envir</code> and the global environment and then reproducibly tracked as dependencies.
<code>verbose</code>	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing.

	<ul style="list-style-type: none"> • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
hook	Deprecated.
cache	drake cache as created by <code>new_cache()</code> . See also <code>drake_cache()</code> .
fetch_cache	Deprecated.
parallelism	<p>Character scalar, type of parallelism to use. For detailed explanations, see the high-performance computing chapter # nolint of the user manual.</p> <p>You could also supply your own scheduler function if you want to experiment or aggressively optimize. The function should take a single config argument (produced by <code>drake_config()</code>). Existing examples from drake's internals are the <code>backend_*</code>() functions:</p> <ul style="list-style-type: none"> • <code>backend_loop()</code> • <code>backend_clustermq()</code> • <code>backend_future()</code> However, this functionality is really a back door and should not be used for production purposes unless you really know what you are doing and you are willing to suffer setbacks whenever drake's un-exported core functions are updated.
jobs	Maximum number of parallel workers for processing the targets. You can experiment with <code>predict_runtime()</code> to help decide on an appropriate number of jobs. For details, visit https://ropenscilabs.github.io/drake-manual/time.html .
jobs_preprocess	Number of parallel jobs for processing the imports and doing other preprocessing tasks.
packages	Character vector packages to load, in the order they should be loaded. Defaults to <code>rev(.packages())</code> , so you should not usually need to set this manually. Just call <code>library()</code> to load your packages before <code>make()</code> . However, sometimes packages need to be strictly forced to load in a certain order, especially if <code>parallelism</code> is "Makefile". To do this, do not use <code>library()</code> or <code>require()</code> or <code>loadNamespace()</code> or <code>attachNamespace()</code> to load any libraries beforehand. Just list your packages in the <code>packages</code> argument in the order you want them to be loaded.
lib_loc	Character vector, optional. Same as in <code>library()</code> or <code>require()</code> . Applies to the <code>packages</code> argument (see above).
prework	Expression (language object), list of expressions, or character vector. Code to run right before targets build. Called only once if <code>parallelism</code> is "loop" and once per target otherwise. This code can be used to set global options, etc.
prepend	Deprecated.
command	Deprecated.
args	Deprecated.
recipe_command	Deprecated.
timeout	deprecated. Use <code>elapsed</code> and <code>cpu</code> instead.
cpu	Same as the <code>cpu</code> argument of <code>setTimeLimit()</code> . Seconds of cpu time before a target times out. Assign target-level cpu timeout times with an optional <code>cpu</code> column in plan.
elapsed	Same as the <code>elapsed</code> argument of <code>setTimeLimit()</code> . Seconds of elapsed time before a target times out. Assign target-level elapsed timeout times with an optional <code>elapsed</code> column in plan.

retries	Number of retries to execute if the target fails. Assign target-level retries with an optional retries column in plan.
force	Logical. If FALSE (default) then drake imposes checks if the cache was created with an old and incompatible version of drake. If there is an incompatibility, make() stops to give you an opportunity to downgrade drake to a compatible version rather than rerun all your targets from scratch.
log_progress	Logical, whether to log the progress of individual targets as they are being built. Progress logging creates extra files in the cache (usually the .drake/ folder) and slows down make() a little. If you need to reduce or limit the number of files in the cache, call make(log_progress = FALSE, recover = FALSE).
graph	An igraph object from the previous make(). Supplying a pre-built graph could save time.
trigger	Name of the trigger to apply to all targets. Ignored if plan has a trigger column. See trigger() for details.
skip_targets	Logical, whether to skip building the targets in plan and just import objects and files.
skip_imports	Logical, whether to totally neglect to process the imports and jump straight to the targets. This can be useful if your imports are massive and you just want to test your project, but it is bad practice for reproducible data analysis. This argument is overridden if you supply your own graph argument.
skip_safety_checks	Logical, whether to skip the safety checks on your workflow. Use at your own peril.
lazy_load	<p>Either a character vector or a logical. Choices:</p> <ul style="list-style-type: none"> • "eager": no lazy loading. The target is loaded right away with assign(). • "promise": lazy loading with delayedAssign() • "bind": lazy loading with active bindings: <code>bindr::populate_env()</code>. • TRUE: same as "promise". • FALSE: same as "eager". <p>lazy_load should not be "promise" for "parLapply" parallelism combined with jobs greater than 1. For local multi-session parallelism and lazy loading, try <code>library(future); future::plan(multisession)</code> and then <code>make(..., parallelism = "future_lapply", lazy_load = "bind")</code>.</p> <p>If lazy_load is "eager", drake prunes the execution environment before each target/stage, removing all superfluous targets and then loading any dependencies it will need for building. In other words, drake prepares the environment in advance and tries to be memory efficient. If lazy_load is "bind" or "promise", drake assigns promises to load any dependencies at the last minute. Lazy loading may be more memory efficient in some use cases, but it may duplicate the loading of dependencies, costing time.</p>
session_info	Logical, whether to save the sessionInfo() to the cache. This behavior is recommended for serious make() s for the sake of reproducibility. This argument only exists to speed up tests. Apparently, sessionInfo() is a bottleneck for small make() s.
cache_log_file	Name of the CSV cache log file to write. If TRUE, the default file name is used (drake_cache.CSV). If NULL, no file is written. If activated, this option writes a flat text file to represent the state of the cache (fingerprints of all the targets and imports). If you put the log file under version control, your commit history will

give you an easy representation of how your results change over time as the rest of your project changes. Hopefully, this is a step in the right direction for data reproducibility.

seed	<p>Integer, the root pseudo-random number generator seed to use for your project. In <code>make()</code>, drake generates a unique local seed for each target using the global seed and the target name. That way, different pseudo-random numbers are generated for different targets, and this pseudo-randomness is reproducible.</p> <p>To ensure reproducibility across different R sessions, <code>set.seed()</code> and <code>.Random.seed</code> are ignored and have no affect on drake workflows. Conversely, <code>make()</code> does not usually change <code>.Random.seed</code>, even when pseudo-random numbers are generated. The exception to this last point is <code>make(parallelism = "clustermq")</code> because the <code>clustermq</code> package needs to generate random numbers to set up ports and sockets for ZeroMQ.</p> <p>On the first call to <code>make()</code> or <code>drake_config()</code>, drake uses the random number generator seed from the <code>seed</code> argument. Here, if the seed is <code>NULL</code> (default), drake uses a seed of 0. On subsequent <code>make()</code>s for existing projects, the project's cached seed will be used in order to ensure reproducibility. Thus, the <code>seed</code> argument must either be <code>NULL</code> or the same seed from the project's cache (usually the <code>.drake/</code> folder). To reset the random number generator seed for a project, use <code>clean(destroy = TRUE)</code>.</p>
caching	<p>Character string, either "master" or "worker".</p> <ul style="list-style-type: none"> "master": Targets are built by remote workers and sent back to the master process. Then, the master process saves them to the cache (<code>config\$cache</code>, usually a file system stor). Appropriate if remote workers do not have access to the file system of the calling R session. Targets are cached one at a time, which may be slow in some situations. "worker": Remote workers not only build the targets, but also save them to the cache. Here, caching happens in parallel. However, remote workers need to have access to the file system of the calling R session. Transferring target data across a network can be slow.
keep_going	Logical, whether to still keep running <code>make()</code> if targets fail.
session	Deprecated. Has no effect now.
pruning_strategy	Deprecated. See <code>memory_strategy</code> .
makefile_path	Deprecated.
console_log_file	<p>Optional character scalar of a file name or connection object (such as <code>stdout()</code>) to dump maximally verbose log information for <code>make()</code> and other functions (all functions that accept a <code>config</code> argument, plus <code>drake_config()</code>). If you choose to use a text file as the console log, it will persist over multiple function calls until you delete it manually. Fields in each row the log file, from left to right: - The node name (short host name) of the computer (from <code>Sys.info()["nodename"]</code>). - The process ID (from <code>Sys.getpid()</code>). - A timestamp with the date and time (in microseconds). - A brief description of what drake was doing. The fields are separated by pipe symbols (<code>" "</code>).</p>
ensure_workers	Deprecated.
garbage_collection	<p>Logical, whether to call <code>gc()</code> each time a target is built during <code>make()</code>.</p>

template	<p>A named list of values to fill in the <code>{{ ... }}</code> placeholders in template files (e.g. from <code>drake_hpc_template_file()</code>). Same as the template argument of <code>clustermq::Q()</code> and <code>clustermq::workers</code>. Enabled for <code>clustermq</code> only (<code>make(parallelism = "clustermq")</code>), not future or batchtools so far. For more information, see the <code>clustermq</code> package: https://github.com/mschubert/clustermq. Some template placeholders such as <code>{{ job_name }}</code> and <code>{{ n_jobs }}</code> cannot be set this way.</p>
sleep	<p>Optional function on a single numeric argument <code>i</code>. Default: <code>function(i) 0.01</code>. To conserve memory, drake assigns a brand new closure to <code>sleep</code>, so your custom function should not depend on in-memory data except from loaded packages.</p> <p>For parallel processing, drake uses a central master process to check what the parallel workers are doing, and for the affected high-performance computing workflows, wait for data to arrive over a network. In between loop iterations, the master process sleeps to avoid throttling. The <code>sleep</code> argument to <code>make()</code> and <code>drake_config()</code> allows you to customize how much time the master process spends sleeping.</p> <p>The <code>sleep</code> argument is a function that takes an argument <code>i</code> and returns a numeric scalar, the number of seconds to supply to <code>Sys.sleep()</code> after iteration <code>i</code> of checking. (Here, <code>i</code> starts at 1.) If the checking loop does something other than sleeping on iteration <code>i</code>, then <code>i</code> is reset back to 1.</p> <p>To sleep for the same amount of time between checks, you might supply something like <code>function(i) 0.01</code>. But to avoid consuming too many resources during heavier and longer workflows, you might use an exponential back-off: say, <code>function(i) { 0.1 + 120 * pexp(i - 1, rate = 0.01) }</code>.</p>
hasty_build	<p>A user-defined function. In "hasty mode" (<code>make(parallelism = "hasty")</code>) this is the function that evaluates a target's command and returns the resulting value. The <code>hasty_build</code> argument has no effect if <code>parallelism</code> is any value other than "hasty".</p> <p>The function you pass to <code>hasty_build</code> must have arguments <code>target</code> and <code>config</code>. Here, <code>target</code> is a character scalar naming the target being built, and <code>config</code> is a configuration list of runtime parameters generated by <code>drake_config()</code>.</p>
memory_strategy	<p>Character scalar, name of the strategy drake uses to load/unload a target's dependencies in memory. You can give each target its own memory strategy, (e.g. <code>drake_plan(x = 1, y = target(f(x), memory_strategy = "lookahead"))</code>) to override the global memory strategy. Choices:</p> <ul style="list-style-type: none"> • "speed": Once a target is newly built or loaded in memory, just keep it there. This choice maximizes speed and hogs memory. • "autoclean": Just before building each new target, unload everything from memory except the target's direct dependencies. After a target is built, discard it from memory. (Set <code>garbage_collection = TRUE</code> to make sure it is really gone.) This option conserves memory, but it sacrifices speed because each new target needs to reload any previously unloaded targets from storage. • "preclean": Just before building each new target, unload everything from memory except the target's direct dependencies. After a target is built, keep it in memory until drake determines they can be unloaded. This option conserves memory, but it sacrifices speed because each new target needs to reload any previously unloaded targets from storage.

- "lookahead": Just before building each new target, search the dependency graph to find targets that will not be needed for the rest of the current `make()` session. After a target is built, keep it in memory until the next memory management stage. In this mode, targets are only in memory if they need to be loaded, and we avoid superfluous reads from the cache. However, searching the graph takes time, and it could even double the computational overhead for large projects.
- "unload": Just before building each new target, unload all targets from memory. After a target is built, **do not** keep it in memory. This mode aggressively optimizes for both memory and speed, but in commands and triggers, you have to manually load any dependencies you need using `readd()`.
- "none": Do not manage memory at all. Do not load or unload anything before building targets. After a target is built, **do not** keep it in memory. This mode aggressively optimizes for both memory and speed, but in commands and triggers, you have to manually load any dependencies you need using `readd()`.

For even more direct control over which targets drake keeps in memory, see the help file examples of `drake_envir()`. Also see the `garbage_collection` argument of `make()` and `drake_config()`.

layout	<code>config\$layout</code> , where <code>config</code> is the return value from a prior call to <code>drake_config()</code> . If your plan or environment have changed since the last <code>make()</code> , do not supply a <code>layout</code> argument. Otherwise, supplying one could save time.
lock_envir	Logical, whether to lock <code>config\$envir</code> during <code>make()</code> . If <code>TRUE</code> , <code>make()</code> quits in error whenever a command in your drake plan (or prework) tries to add, remove, or modify non-hidden variables in your environment/workspace/R session. This is extremely important for ensuring the purity of your functions and the reproducibility/credibility/trust you can place in your project. <code>lock_envir</code> will be set to a default of <code>TRUE</code> in drake version 7.0.0 and higher.
history	Logical, whether to record the build history of your targets. You can also supply a <code>txtq</code> , which is how drake records history. Must be <code>TRUE</code> for <code>drake_history()</code> to work later.
recover	Logical, whether to activate automated data recovery. The default is <code>FALSE</code> because <ol style="list-style-type: none"> 1. Automated data recovery is still experimental. 2. It has reproducibility issues. Targets recovered from the distant past may have been generated with earlier versions of R and earlier package environments that no longer exist.

How it works: if `recover` is `TRUE`, drake tries to salvage old target values from the cache instead of running commands from the plan. A target is recoverable if

1. There is an old value somewhere in the cache that shares the command, dependencies, etc. of the target about to be built.
2. The old value was generated with `make(recoverable = TRUE)`.

If both conditions are met, drake will

1. Assign the most recently-generated admissible data to the target, and
2. skip the target's command.

Functions `recoverable()` and `r_recoverable()` show the most upstream outdated targets that will be recovered in this way in the next `make()` or `r_make()`.

recoverable	Logical, whether to make target values recoverable with <code>make(recover = TRUE)</code> . This requires writing extra files to the cache, and it prevents old metadata from being removed with garbage collection (<code>clean(garbage_collection = TRUE)</code> , <code>gc()</code> in storrs). If you need to limit the cache size or the number of files in the cache, consider <code>make(recoverable = FALSE, progress = FALSE)</code> .
curl_handles	A named list of curl handles. Each value is an object from <code>curl::new_handle()</code> , and each name is a URL (and should start with "http", "https", or "ftp"). Example: <code>list(http://httpbin.org/basic-auth = curl::new_handle(username = "user", password = "passwd"))</code> Then, if your plan has <code>file_in("http://httpbin.org/basic-auth/...")</code> , drake will authenticate using the username and password of the handle for <code>http://httpbin.org/basic-auth/...</code> . drake uses partial matching on text to find the right handle of the <code>file_in()</code> URL, so the name of the handle could be the complete URL (" <code>http://httpbin.org/basic-auth/...</code> ") or a part of the URL (e.g. " <code>http://httpbin.org/</code> " or " <code>http://httpbin.org/basic-auth/</code> "). If you have multiple handles whose names match your URL, drake will choose the closest match.

Details

Once you create a list with `drake_config()`, do not modify it by hand.

Utility functions such as `outdated()`, `vis_drake_graph()`, and `predict_runtime()` require output from `drake_config()` for the `config` argument. If you supply a `drake_config()` object to the `config` argument of `make()`, then drake will ignore all the other arguments because it already has everything it needs in `config`.

Value

The master internal configuration list of a project.

Recovery

`make(recover = TRUE, recoverable = TRUE)` powers automated data recovery. The default of `recover` is `FALSE` because

1. Automated data recovery is still experimental.
2. It has reproducibility issues. Targets recovered from the distant past may have been generated with earlier versions of R and earlier package environments that no longer exist.

How it works: if `recover` is `TRUE`, drake tries to salvage old target values from the cache instead of running commands from the plan. A target is recoverable if

1. There is an old value somewhere in the cache that shares the command, dependencies, etc. of the target about to be built.
2. The old value was generated with `make(recoverable = TRUE)`.

If both conditions are met, drake will

1. Assign the most recently-generated admissible data to the target, and
2. skip the target's command.

See Also

`make()`, `drake_plan()`, `vis_drake_graph()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  load_mtcars_example() # Get the code with drake_example("mtcars").
  # Construct the master internal configuration list.
  config <- drake_config(my_plan)
  if (requireNamespace("visNetwork")) {
    vis_drake_graph(config) # See the dependency graph.
    if (requireNamespace("networkD3")) {
      sankey_drake_graph(config) # See the dependency graph.
    }
  }
  # These functions are faster than otherwise
  # because they use the configuration list.
  outdated(config) # Which targets are out of date?
  missed(config) # Which imports are missing?
})

## End(Not run)
```

drake_debug

Run a single target's command in debug mode. **Maturing**

Description

`drake_debug()` loads a target's dependencies and then runs its command in debug mode (see `browser()`, `debug()`, and `debugonce()`). This function does not store the target's value in the cache (see <https://github.com/ropensci/drake/issues/587>).

Usage

```
drake_debug(target = NULL, config = NULL, character_only = FALSE,
  envir = NULL, jobs = 1, replace = FALSE, verbose = TRUE)
```

Arguments

<code>target</code>	Name of the target.
<code>config</code>	Internal configuration list.
<code>character_only</code>	Logical, whether name should be treated as a character or a symbol (just like <code>character.only</code> in <code>library()</code>).
<code>envir</code>	Environment to load objects into. Defaults to the calling environment (current workspace).
<code>jobs</code>	Number of parallel jobs for loading objects. On non-Windows systems, the loading process for multiple objects can be lightly parallelized via <code>parallel::mclapply()</code> . just set <code>jobs</code> to be an integer greater than 1. On Windows, <code>jobs</code> is automatically demoted to 1.
<code>replace</code>	Logical. If <code>FALSE</code> , items already in your environment will not be replaced.
<code>verbose</code>	Logical, whether to print out the target you are debugging.

Value

The value of the target right after it is built.

See Also

`drake_build()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    # This example is not really a user-side demonstration.
    # It just walks through a dive into the internals.
    # Populate your workspace and write 'report.Rmd'.
    load_mtcars_example() # Get the code with drake_example("mtcars").
    # Create the master internal configuration list.
    config <- drake_config(my_plan)
    out <- drake_build(small, config = config)
    # Now includes `small`.
    cached()
    head(readadd(small))
    # `small` was invisibly returned.
    head(out)
    # If you previously called make(),
    # `config` is just read from the cache.
    make(my_plan, verbose = FALSE)
    result <- drake_build(small, config = config)
    head(result)
  }
})

## End(Not run)
```

drake_envir

Get the environment where drake builds targets **Questioning**

Description

Call this function inside the commands in your plan to get the environment where drake builds targets. That way, you can strategically remove targets from memory while `make()` is running. That way, you can limit the amount of computer memory you use.

Usage

```
drake_envir()
```

Value

The environment where drake builds targets.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`from_plan()`

Examples

```
## Not run:
isolate_example("contain side effects", {
  plan <- drake_plan(
    large_data_1 = sample.int(1e4),
    large_data_2 = sample.int(1e4),
    subset = c(large_data_1[seq_len(10)], large_data_2[seq_len(10)]),
    summary = {
      print(ls(envir = drake_envir()))
      # We don't need the large_data_* targets in memory anymore.
      rm(large_data_1, large_data_2, envir = drake_envir())
      print(ls(envir = drake_envir()))
      mean(subset)
    }
  )
  make(plan, cache = storrr::storrr_environment(), session_info = FALSE)
})

## End(Not run)
```

drake_example

*Download the files of an example drake project. **Maturing*****Description**

The `drake_example()` function downloads a folder from <https://github.com/wlandau/drake-examples>. By default, it creates a new folder with the example name in your current working directory. After the files are written, have a look at the enclosed README file. Other instructions are available in the files at <https://github.com/wlandau/drake-examples>.

Usage

```
drake_example(example = "main", to = getwd(), destination = NULL,
              overwrite = FALSE, quiet = TRUE)
```

Arguments

<code>example</code>	Name of the example. The possible values are the names of the folders at https://github.com/wlandau/drake-examples .
<code>to</code>	Character scalar, the folder containing the code files for the example. passed to the <code>exdir</code> argument of <code>utils::unzip()</code> .
<code>destination</code>	Deprecated; use <code>to</code> instead.
<code>overwrite</code>	Logical, whether to overwrite an existing folder with the same name as the drake example.
<code>quiet</code>	Logical, passed to <code>download::download()</code> and thus <code>utils::download.file()</code> . Whether to download quietly or print progress.

Value

NULL

See Also

[drake_examples\(\)](#), [make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (requireNamespace("downloader")) {
    drake_examples() # List all the drake examples.
    # Sets up the same example from load_mtcars_example()
    drake_example("mtcars")
    # Sets up the SLURM example.
    drake_example("slurm")
  }
})

## End(Not run)
```

drake_examples	<i>List the names of all the drake examples.</i> Stable
----------------	--

Description

You can find the code files of the examples at <https://github.com/wlandau/drake-examples>. The `drake_examples()` function downloads the list of examples from <https://wlandau.github.io/drake-examples/examples.md>, so you need an internet connection.

Usage

```
drake_examples(quiet = TRUE)
```

Arguments

quiet	Logical, passed to <code>downloader::download()</code> and thus <code>utils::download.file()</code> . Whether to download quietly or print progress.
-------	--

Value

Names of all the drake examples.

See Also

[drake_example\(\)](#), [make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (requireNamespace("downloader")) {
    drake_examples() # List all the drake examples.
    # Sets up the example from load_mtcars_example()
    drake_example("mtcars")
    # Sets up the SLURM example.
    drake_example("slurm")
  }
})

## End(Not run)
```

drake_gc	<i>Do garbage collection on the drake cache.</i> Maturing
----------	--

Description

Garbage collection removes obsolete target values from the cache.

Usage

```
drake_gc(path = NULL, search = NULL, verbose = 1L,
         cache = drake::drake_cache(path = path, verbose = verbose),
         force = FALSE)
```

Arguments

path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
search	Deprecated.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
cache	drake cache. See new_cache() . If supplied, path is ignored.
force	Logical, whether to load the cache despite any back compatibility issues with the running version of drake.

Details

Caution: garbage collection *actually* removes data so it is no longer recoverable with [drake_history\(\)](#) or `make(recover = TRUE)`. You cannot undo this operation. Use at your own risk.

Value

`NULL`

See Also

[clean\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    # At this point, check the size of the '.drake/' cache folder.
    # Clean without garbage collection.
    clean(garbage_collection = FALSE)
    # The '.drake/' cache folder is still about the same size.
    drake_gc() # Do garbage collection on the cache.
    # The '.drake/' cache folder should have gotten much smaller.
  }
})

## End(Not run)
```

`drake_get_session_info`*Session info of the last call to `make()`. **Maturing***

Description

By default, session info is saved during `make()` to ensure reproducibility. Your loaded packages and their versions are recorded, for example.

Usage

```
drake_get_session_info(path = NULL, search = NULL,
  cache = drake::drake_cache(path = path, verbose = verbose),
  verbose = 1L)
```

Arguments

<code>path</code>	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
<code>search</code>	Deprecated.
<code>cache</code>	drake cache. See <code>new_cache()</code> . If supplied, <code>path</code> is ignored.
<code>verbose</code>	Integer, control printing to the console/terminal. <ul style="list-style-type: none">• 0: print nothing.• 1: print targets, retries, and failures.• 2: also show a spinner when preprocessing tasks are underway.

Value

`sessionInfo()` of the last call to `make()`

See Also

`diagnose()`, `cached()`, `readd()`, `drake_plan()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    drake_get_session_info() # Get the cached sessionInfo() of the last make().
  }
})

## End(Not run)
```

drake_ggraph

Visualize the workflow with ggraph/ggplot2 **Stable****Description**

This function requires packages ggplot2 and ggraph. Install them with `install.packages(c("ggplot2", "ggraph"))`

Usage

```
drake_ggraph(config, build_times = "build", digits = 3,
  targets_only = FALSE, main = NULL, from = NULL, mode = c("out",
    "in", "all"), order = NULL, subset = NULL, make_imports = TRUE,
    from_scratch = FALSE, full_legend = FALSE, group = NULL,
    clusters = NULL, show_output_files = TRUE, label_nodes = FALSE,
    transparency = TRUE)
```

Arguments

config	A drake_config() configuration list. You can get one as a return value from make() as well.
build_times	Character string or logical. If character, the choices are 1. "build": runtime of the command plus the time it take to store the target or import. 2. "command": just the runtime of the command. 3. "none": no build times. If logical, build_times selects whether to show the times from 'build_times(..., type = "build")' or use no build times at all. See build_times() for details.
digits	Number of digits for rounding the build times
targets_only	Logical, whether to skip the imports and only include the targets in the workflow plan.
main	Character string, title of the graph.
from	Optional collection of target/import names. If from is nonempty, the graph will restrict itself to a neighborhood of from. Control the neighborhood with mode and order.
mode	Which direction to branch out in the graph to create a neighborhood around from. Use "in" to go upstream, "out" to go downstream, and "all" to go both ways and disregard edge direction altogether.
order	How far to branch out to create a neighborhood around from. Defaults to as far as possible. If a target is in the neighborhood, then so are all of its custom file_out() files if show_output_files is TRUE. That means the actual graph order may be slightly greater than you might expect, but this ensures consistency between show_output_files = TRUE and show_output_files = FALSE.
subset	Optional character vector. Subset of targets/imports to display in the graph. Applied after from, mode, and order. Be advised: edges are only kept for adjacent nodes in subset. If you do not select all the intermediate nodes, edges will drop from the graph.
make_imports	Logical, whether to make the imports first. Set to FALSE to increase speed and risk using obsolete information.
from_scratch	Logical, whether to assume all the targets will be made from scratch on the next make() . Makes all targets outdated, but keeps information about build progress in previous make() s.

full_legend	Logical. If TRUE, all the node types are printed in the legend. If FALSE, only the node types used are printed in the legend.
group	Optional character scalar, name of the column used to group nodes into columns. All the columns names of your original drake plan are choices. The other choices (such as "status") are column names in the nodes . To group nodes into clusters in the graph, you must also supply the clusters argument.
clusters	Optional character vector of values to cluster on. These values must be elements of the column of the nodes data frame that you specify in the group argument to drake_graph_info().
show_output_files	Logical, whether to include <code>file_out()</code> files in the graph.
label_nodes	Logical, whether to label the nodes. If FALSE, the graph will not have any text next to the nodes, which is recommended for large graphs with lots of targets.
transparency	Logical, whether to allow transparency in the rendered graph. Set to FALSE if you get warnings like "semi-transparency is not supported on this device".

Value

A ggplot2 object, which you can modify with more layers, show with `plot()`, or save as a file with `ggsave()`.

See Also

[vis_drake_graph\(\)](#), [sankey_drake_graph\(\)](#), [render_drake_ggraph\(\)](#), [text_drake_graph\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  load_mtcars_example() # Get the code with drake_example("mtcars").
  config <- drake_config(my_plan)
  # Plot the network graph representation of the workflow.
  if (requireNamespace("ggraph", quietly = TRUE)) {
    drake_ggraph(config) # Save to a file with `ggplot2::ggsave()`.
  }
})

## End(Not run)
```

drake_graph_info

Prepare the workflow graph for visualization **Stable**

Description

With the returned data frames, you can plot your own custom `visNetwork` graph.

Usage

```
drake_graph_info(config, from = NULL, mode = c("out", "in", "all"),
  order = NULL, subset = NULL, build_times = "build", digits = 3,
  targets_only = FALSE, font_size = 20, from_scratch = FALSE,
  make_imports = TRUE, full_legend = FALSE, group = NULL,
  clusters = NULL, show_output_files = TRUE, hover = FALSE)
```

Arguments

config	A drake_config() configuration list. You can get one as a return value from make() as well.
from	Optional collection of target/import names. If from is nonempty, the graph will restrict itself to a neighborhood of from. Control the neighborhood with mode and order.
mode	Which direction to branch out in the graph to create a neighborhood around from. Use "in" to go upstream, "out" to go downstream, and "all" to go both ways and disregard edge direction altogether.
order	How far to branch out to create a neighborhood around from. Defaults to as far as possible. If a target is in the neighborhood, then so are all of its custom file_out() files if show_output_files is TRUE. That means the actual graph order may be slightly greater than you might expect, but this ensures consistency between show_output_files = TRUE and show_output_files = FALSE.
subset	Optional character vector. Subset of targets/imports to display in the graph. Applied after from, mode, and order. Be advised: edges are only kept for adjacent nodes in subset. If you do not select all the intermediate nodes, edges will drop from the graph.
build_times	Character string or logical. If character, the choices are 1. "build": runtime of the command plus the time it take to store the target or import. 2. "command": just the runtime of the command. 3. "none": no build times. If logical, build_times selects whether to show the times from 'build_times(..., type = "build")' or use no build times at all. See build_times() for details.
digits	Number of digits for rounding the build times
targets_only	Logical, whether to skip the imports and only include the targets in the workflow plan.
font_size	Numeric, font size of the node labels in the graph
from_scratch	Logical, whether to assume all the targets will be made from scratch on the next make() . Makes all targets outdated, but keeps information about build progress in previous make() s.
make_imports	Logical, whether to make the imports first. Set to FALSE to increase speed and risk using obsolete information.
full_legend	Logical. If TRUE, all the node types are printed in the legend. If FALSE, only the node types used are printed in the legend.
group	Optional character scalar, name of the column used to group nodes into columns. All the columns names of your original drake plan are choices. The other choices (such as "status") are column names in the nodes . To group nodes into clusters in the graph, you must also supply the clusters argument.
clusters	Optional character vector of values to cluster on. These values must be elements of the column of the nodes data frame that you specify in the group argument to drake_graph_info() .
show_output_files	Logical, whether to include file_out() files in the graph.
hover	Logical, whether to show text (file contents, commands, etc.) when you hover your cursor over a node.

Value

A list of three data frames: one for nodes, one for edges, and one for the legend nodes. The list also contains the default title of the graph.

See Also

[vis_drake_graph\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (requireNamespace("visNetwork", quietly = TRUE)) {
    if (suppressWarnings(require("knitr"))){
      load_mtcars_example() # Get the code with drake_example("mtcars").
      config <- drake_config(my_plan) # my_plan loaded with load_mtcars_example()
      vis_drake_graph(config) # Jump straight to the interactive graph.
      # Get a list of data frames representing the nodes, edges,
      # and legend nodes of the visNetwork graph from vis_drake_graph().
      raw_graph <- drake_graph_info(config = config)
      # Choose a subset of the graph.
      smaller_raw_graph <- drake_graph_info(
        config = config,
        from = c("small", "reg2"),
        mode = "in"
      )
      # Inspect the raw graph.
      str(raw_graph)
      # Use the data frames to plot your own custom visNetwork graph.
      # For example, you can omit the legend nodes
      # and change the direction of the graph.
      library(visNetwork)
      graph <- visNetwork(nodes = raw_graph$nodes, edges = raw_graph$edges)
      visHierarchicalLayout(graph, direction = 'UD')
    }
  }
})

## End(Not run)
```

drake_history

History and provenance **Experimental**

Description

See the history and provenance of your targets: what you ran, when you ran it, the function arguments you used, and how to get old data back.

Usage

```
drake_history(cache = NULL, history = NULL, analyze = TRUE,
  verbose = TRUE)
```

Arguments

cache	drake cache as created by <code>new_cache()</code> . See also <code>drake_cache()</code> .
history	Logical, whether to record the build history of your targets. You can also supply a <code>txtq</code> , which is how drake records history. Must be TRUE for <code>drake_history()</code> to work later.
analyze	Logical, whether to analyze <code>drake_plan()</code> commands for arguments to function calls. Could be slow because this requires parsing and analyzing lots of R code.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.

Details

`drake_history()` returns a data frame with the following columns.

- target: the name of the target.
- current: logical, whether the row describes the data actually assigned to the target name in the cache, e.g. what you get with `load(target)` and `readd(target)`. Does **NOT** tell you if the target is up to date.
- built: when the target's value was stored in the cache. This is the true creation date of the target's value, not the recovery date from `make(recover = TRUE)`.
- exists: logical, whether the target's historical value still exists in the cache. Garbage collection via `(clean(garbage_collection = TRUE) and drake_cache()$gc())` remove these historical values, but `clean()` under the default settings does not.
- hash: fingerprint of the target's historical value in the cache. If the value still exists, you can read it with `drake_cache()$get_value(hash)`.
- command: the `drake_plan()` command executed to build the target.
- seed: random number generator seed.
- runtime: the time it took to execute the `drake_plan()` command. Does not include overhead due to drake's processing.

If `analyze` is TRUE, various other columns are included to show the explicitly-named length-1 arguments to function calls in the commands. See the "Provenance" section for more details.

Value

A data frame of target history.

Provenance

If `analyze` is TRUE, drake scans your `drake_plan()` commands for function arguments and mentions them in the history. A function argument shows up if and only if 1. It has length 1. 2. It is atomic, i.e. a base type: logical, integer, real, complex, character, or raw. 3. It is explicitly named in the function call, For example, `x` is detected as 1 in `fn(list(x = 1))` but not `f(list(1))`. The exceptions are `file_out()`, `file_in()`, and `knitr_in()`. For example, `filename` is detected as "my_file.csv" in `process_data(filename = file_in("my_file.csv"))`. NB: in `process_data(filename = file_in("a", "b"))` `filename` is not detected because the value must be atomic.

Examples

```
## Not run:
isolate_example("contain side-effects", {
  # First, let's iterate on a drake workflow.
  load_mtcars_example()
  make(my_plan, history = TRUE, verbose = 0L)
  # Naturally, we'll make updates to our targets along the way.
  reg2 <- function(d) {
    d$x2 <- d$x ^ 3
    lm(y ~ x2, data = d)
  }
  Sys.sleep(0.01)
  make(my_plan, history = TRUE, verbose = 0L)
  # The history is a data frame about all the recorded runs of your targets.
  out <- drake_history(analyze = TRUE)
  print(out)
  # Let's use the history to recover the oldest version
  # of our regression2_small target.
  oldest_reg2_small <- max(which(out$target == "regression2_small"))
  hash_oldest_reg2_small <- out[oldest_reg2_small, ]$hash
  cache <- drake_cache()
  cache$get_value(hash_oldest_reg2_small)
  # If you run clean(), drake can still find all the targets.
  clean(small)
  drake_history()
  # But if you run clean() with garbage collection,
  # older versions of your targets may be gone.
  clean(large, garbage_collection = TRUE)
  drake_history()
  invisible()
})

## End(Not run)
```

drake_hpc_template_file

Write a template file for deploying work to a cluster / job scheduler.

Stable

Description

See the example files from [drake_examples\(\)](#) and [drake_example\(\)](#) for example usage.

Usage

```
drake_hpc_template_file(file = drake::drake_hpc_template_files(),
  to = getwd(), overwrite = FALSE)
```

Arguments

file	Name of the template file, including the "tmpl" extension.
to	Character vector, where to write the file.
overwrite	Logical, whether to overwrite an existing file of the same name.

Value

NULL is returned, but a batchtools template file is written.

See Also

[drake_hpc_template_files\(\)](#), [drake_examples\(\)](#), [drake_example\(\)](#), [shell_file\(\)](#)

Examples

```
## Not run:
plan <- drake_plan(x = rnorm(1e7), y = rnorm(1e7))
# List the available template files.
drake_hpc_template_files()
# Write a SLURM template file.
out <- file.path(tempdir(), "slurm_batchtools.tpl")
drake_hpc_template_file("slurm_batchtools.tpl", to = tempdir())
cat(readLines(out), sep = "\n")
# library(future.batchtools) # nolint
# future::plan(batchtools_slurm, template = out) # nolint
# make(plan, parallelism = "future", jobs = 2) # nolint

## End(Not run)
```

drake_hpc_template_files

*List the available example template files for deploying work to a cluster / job scheduler. **Stable***

Description

See the example files from [drake_examples\(\)](#) and [drake_example\(\)](#) for example usage.

Usage

```
drake_hpc_template_files()
```

Value

A character vector of example template files that you can write with [drake_hpc_template_file\(\)](#).

See Also

[drake_hpc_template_file\(\)](#), [drake_examples\(\)](#), [drake_example\(\)](#), [shell_file\(\)](#)

Examples

```
## Not run:
plan <- drake_plan(x = rnorm(1e7), y = rnorm(1e7))
# List the available template files.
drake_hpc_template_files()
# Write a SLURM template file.
out <- file.path(tempdir(), "slurm_batchtools.tpl")
drake_hpc_template_file("slurm_batchtools.tpl", to = tempdir())
```

```

cat(readLines(out), sep = "\n")
# library(future.batchtools) # nolint
# future::plan(batchtools_slurm, template = out) # nolint
# make(plan, parallelism = "future", jobs = 2) # nolint

## End(Not run)

```

drake_plan	Create a workflow plan data frame for the plan argument of <code>make()</code> . Maturing
------------	---

Description

A drake plan is a data frame with columns "target" and "command". Each target is an R object produced in your workflow, and each command is the R code to produce it.

Usage

```

drake_plan(..., list = character(0), file_targets = NULL,
  strings_in_dots = NULL, tidy_evaluation = NULL, transform = TRUE,
  trace = FALSE, envir = parent.frame(), tidy_eval = TRUE,
  max_expand = NULL)

```

Arguments

...	A collection of symbols/targets with commands assigned to them. See the examples for details.
list	Deprecated
file_targets	Deprecated.
strings_in_dots	Deprecated.
tidy_evaluation	Deprecated. Use tidy_eval instead.
transform	Logical, whether to transform the plan into a larger plan with more targets. Requires the transform field in target(). See the examples for details.
trace	Logical, whether to add columns to show what happens during target transformations.
envir	Environment for tidy evaluation.
tidy_eval	Logical, whether to use tidy evaluation (e.g. unquoting/!!) when resolving commands. Tidy evaluation in transformations is always turned on regardless of the value you supply to this argument.
max_expand	Positive integer, optional upper bound on the lengths of grouping variables for map() and cross(). Comes in handy when you have a massive number of targets and you want to test on a miniature version of your workflow before you scale up to production.

Details

Besides "target" and "command", `drake_plan()` understands a special set of optional columns.

For details, visit <https://ropenscilabs.github.io/drake-manual/plans.html#special-custom-columns-in-y>
nolint

Value

A data frame of targets, commands, and optional custom columns.

Columns

`drake_plan()` creates a special data frame. At minimum, that data frame must have columns `target` and `command` with the target names and the R code chunks to build them, respectively.

You can add custom columns yourself, either with `target()` (e.g. `drake_plan(y = target(f(x), transform = map(c(1, 2)), format = "fst"))`) or by appending columns post-hoc (e.g. `plan$col <- vals`).

Some of these custom columns are special. They are optional, but drake looks for them at various points in the workflow.

- `transform`: a call to `map()`, `split()`, `cross()`, or `combine()` to create and manipulate large collections of targets. Details: (<https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>). # nolint
- `format`: set a storage format to save big targets more efficiently. Most formats are faster than ordinary storage, and they consume far less memory. Available formats:
 - `"fst"`: save big data frames fast. The target must be a data frame, and you must have the `fst` package installed.
 - `"keras"`: save Keras models as HDF5 files. Requires the `keras` package.
 - `"rds"`: save any object. This is similar to the default storage except we avoid creating a serialized copy of the entire target in memory. Requires R \geq 3.5.0 so drake can use ALTREP.
- `trigger`: rule to decide whether a target needs to run. It is recommended that you define this one with `target()`. Details: <https://ropenscilabs.github.io/drake-manual/triggers.html>.
- `hpc`: logical values (TRUE/FALSE/NA) whether to send each target to parallel workers. Visit <https://ropenscilabs.github.io/drake-manual/hpc.html#selectivity> to learn more.
- `resources`: target-specific lists of resources for a computing cluster. See <https://ropenscilabs.github.io/drake-manual/hpc.html#advanced-options> for details.
- `elapsed` and `cpu`: number of seconds to wait for the target to build before timing out (`elapsed` for elapsed time and `cpu` for CPU time).
- `retries`: number of times to retry building a target in the event of an error.
- `seed`: an optional pseudo-random number generator (RNG) seed for each target. drake usually comes up with its own unique reproducible target-specific seeds using the global seed (the `seed` argument to `make()` and `drake_config()`) and the target names, but you can overwrite these automatic seeds. NA entries default back to drake's automatic seeds.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), # nolint supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), # nolint or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.

- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

DSL

drake has special syntax for generating large plans. Your code will look something like `drake_plan(x = target(cmd, transform = f(y, z), group = g)` where `f()` is either `map()`, `cross()`, `split()`, or `combine()` (similar to `purrr::pmap()`, `tidy::crossing()`, `base::split()`, and `dplyr::summarize()`, respectively). These verbs mimic Tidyverse behavior to scale up existing plans to large numbers of targets. You can read about this interface at <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>. `# nolint`

See Also

`make`, `drake_config`, `transform_plan`, `map`, `split`, `cross`, `combine`

Examples

```
## Not run:
isolate_example("contain side effects", {
  # For more examples, visit
  # https://ropenscilabs.github.io/drake-manual/plans.html.

  # Create drake plans:
  mtcars_plan <- drake_plan(
    write.csv(mtcars[, c("mpg", "cyl")], file_out("mtcars.csv")),
    value = read.csv(file_in("mtcars.csv"))
  )
  mtcars_plan
  make(mtcars_plan) # Makes `mtcars.csv` and then `value`
  head(readd(value))
  # You can use knitr inputs too. See the top command below.

  load_mtcars_example()
  head(my_plan)

  # The `knitr_in("report.Rmd")` tells `drake` to dive into the active
  # code chunks to find dependencies.
  # There, `drake` sees that `small`, `large`, and `coef_regression2_small`
  # are loaded in with calls to `load()` and `readd()`.
  deps_code("report.Rmd")

  # Formats are great for big data: https://github.com/ropensci/drake/pull/977
```

```

# Below, each target is 1.6 GB in memory.
# Run make() on this plan to see how much faster fst is!
n <- 1e8
plan <- drake_plan(
  data_fst = target(
    data.frame(x = runif(n), y = runif(n)),
    format = "fst"
  ),
  data_old = data.frame(x = runif(n), y = runif(n))
)

# Use transformations to generate large plans.
# Read more at
# <https://ropenscilabs.github.io/drake-manual/plans.html#create-large-plans-the-easy-way>. # nolint
drake_plan(
  data = target(
    simulate(nrows),
    transform = map(nrows = c(48, 64)),
    custom_column = 123
  ),
  reg = target(
    reg_fun(data),
    transform = cross(reg_fun = c(reg1, reg2), data)
  ),
  summ = target(
    sum_fun(data, reg),
    transform = cross(sum_fun = c(coef, residuals), reg)
  ),
  winners = target(
    min(summ),
    transform = combine(summ, .by = c(data, sum_fun))
  )
)

# Split data among multiple targets.
drake_plan(
  large_data = get_data(),
  slice_analysis = target(
    large_data %>%
      analyze(),
    transform = split(large_data, slices = 4)
  ),
  results = target(
    rbind(slice_analysis),
    transform = combine(slice_analysis)
  )
)

# Set trace = TRUE to show what happened during the transformation process.
drake_plan(
  data = target(
    simulate(nrows),
    transform = map(nrows = c(48, 64)),
    custom_column = 123
  ),
  reg = target(
    reg_fun(data),

```



```

    transform = cross(reg_fun = c(reg1, reg2), data)
  ),
  summ = target(
    sum_fun(data, reg),
    transform = cross(sum_fun = c(coef, residuals), reg)
  ),
  winners = target(
    min(summ),
    transform = combine(summ, .by = c(data, sum_fun))
  ),
  trace = TRUE
)

# You can create your own custom columns too.
# See ?triggers for more on triggers.
drake_plan(
  website_data = target(
    command = download_data("www.your_url.com"),
    trigger = "always",
    custom_column = 5
  ),
  analysis = analyze(website_data)
)

# Tidy evaluation can help generate super large plans.
sms <- rlang::syms(letters) # To sub in character args, skip this.
drake_plan(x = target(f(char), transform = map(char = !!sms)))
})

## End(Not run)

```

drake_plan_source

Show the code required to produce a given drake plan **Stable**

Description

You supply a plan, and `drake_plan_source()` supplies code to generate that plan. If you have the [prettycode package](#), installed, you also get nice syntax highlighting in the console when you print it.

Usage

```
drake_plan_source(plan)
```

Arguments

`plan` A workflow plan data frame (see `drake_plan()`)

Value

a character vector of lines of text. This text is a call to `drake_plan()` that produces the plan you provide.

See Also[drake_plan\(\)](#)**Examples**

```

plan <- drake::drake_plan(
  small_data = download_data("https://some_website.com"),
  large_data_raw = target(
    command = download_data("https://lots_of_data.com"),
    trigger = trigger(
      change = time_last_modified("https://lots_of_data.com"),
      command = FALSE,
      depend = FALSE
    ),
    timeout = 1e3
  )
)
print(plan)
if (requireNamespace("styler", quietly = TRUE)) {
  source <- drake_plan_source(plan)
  print(source) # Install the prettycode package for syntax highlighting.
  file <- tempfile() # Path to an R script to contain the drake_plan() call.
  writeLines(source, file) # Save the code to an R script.
}

```

drake_slice

*Take a strategic subset of a dataset. **Maturing*****Description**

`drake_slice()` is similar to `split()`. Both functions partition data into disjoint subsets, but whereas `split()` returns *all* the subsets, `drake_slice()` returns just *one*. In other words, `drake_slice(..., index = i)` returns `split(...)[[i]]`. Other features: 1. `drake_slice()` works on vectors, data frames, matrices, lists, and arbitrary arrays. 2. Like `parallel::splitIndices()`, `drake_slice()` tries to distribute the data uniformly across subsets. See the examples to learn why splitting is useful in drake.

Usage

```
drake_slice(data, slices, index, margin = 1L, drop = FALSE)
```

Arguments

<code>data</code>	A list, vector, data frame, matrix, or arbitrary array. Anything with a <code>length()</code> or <code>dim()</code> .
<code>slices</code>	Integer of length 1, number of slices (i.e. pieces) of the whole dataset. Remember, <code>drake_slice(index = i)</code> returns only slice number <code>i</code> .
<code>index</code>	Integer of length 1, which piece of the partition to return.
<code>margin</code>	Integer of length 1, margin over which to split the data. For example, for a data frame or matrix, use <code>margin = 1</code> to split over rows and <code>margin = 2</code> to split over columns. Similar to <code>MARGIN</code> in <code>apply()</code> .
<code>drop</code>	Logical, for matrices and arrays. If <code>TRUE</code> , the result is coerced to the lowest possible dimension. See <code>?['</code> for details.

Value

A subset of data.

Examples

```
# Simple usage
x <- matrix(seq_len(20), nrow = 5)
x
drake_slice(x, slices = 3, index = 1)
drake_slice(x, slices = 3, index = 2)
drake_slice(x, slices = 3, index = 3)
drake_slice(x, slices = 3, margin = 2, index = 1)
# In drake, you can split a large dataset over multiple targets.
## Not run:
isolate_example("contain side effects", {
  plan <- drake_plan(
    large_data = iris,
    data_split = target(
      drake_slice(large_data, slices = 50, index = i),
      transform = map(i = !!seq_len(50))
    )
  )
  plan
  cache <- storr::storr_environment()
  make(plan, cache = cache, session_info = FALSE, verbose = FALSE)
  readd(data_split_1L, cache = cache)
  readd(data_split_2L, cache = cache)
})

## End(Not run)
```

expose_imports	<i>Expose all the objects in a package so make() can detect them as imported dependencies.</i> Stable
----------------	--

Description

When drake analyzes the functions in your environment, it understands that some of your functions are nested inside other functions. It dives into nested function after nested function in your environment so that if an inner function changes, targets produced by the outer functions will become out of date. However, drake stops searching as soon as it sees a function from a package. This keeps projects from being too brittle, but it is sometimes problematic. You may want to strongly depend on a package's internals. In fact, you may want to wrap your data analysis project itself in a formal R package, so you want all your functions to be reproducibly tracked.

To make all a package's functions available to be tracked as dependencies, use the `expose_imports()` function. See the examples in this help file for a demonstration.

Usage

```
expose_imports(package, character_only = FALSE, envir = parent.frame(),
  jobs = 1)
```

Arguments

package	Name of the package, either a symbol or a string, depending on character_only.
character_only	Logical, whether to interpret package as a character string or a symbol (quoted vs unquoted).
envir	Environment to load the exposed package imports. You will later pass this envir to make() .
jobs	Number of parallel jobs for the parallel processing of the imports.

Details

Thanks to [Jasper Clarkberg](#) for the idea that makes this function work.

Value

The environment that the exposed imports are loaded into. Defaults to your R workspace.

Examples

```
## Not run:
isolate_example("contain side effects", {
  # Suppose you have a workflow that uses the `digest()` function,
  # which computes the hash of an object.

  library(digest) # Has the digest() function.
  g <- function(x) {
    digest(x)
  }
  f <- function(x) {
    g(x)
  }
  plan <- drake_plan(x = f(1))
  cache <- storr::storr_environment() # optional

  # Here are the reproducibly tracked objects in the workflow.
  config <- drake_config(plan, cache = cache, history = FALSE)
  tracked(config)

  # But the digest() function has dependencies too.
  head(deps_code(digest))

  # Why doesn't `drake` import them? Because it knows `digest()`
  # is from a package, and it doesn't usually dive into functions
  # from packages. We need to call expose_imports() to expose
  # a package's inner functions.

  expose_imports(digest)
  config <- drake_config(plan, cache = cache, history = FALSE)
  new_objects <- tracked(config)
  head(new_objects, 10)
  length(new_objects)

  # Now when you call `make()`, `drake` will dive into `digest`
  # to import dependencies.

  make(plan, cache = cache, history = FALSE)
```

```

head(cached(cache = cache), 10)
length(cached(cache = cache))

# Why would you want to expose a whole package like this?
# Because you may want to wrap up your data science project
# as a formal R package. In that case, `expose_imports()`
# tells `drake` to reproducibly track all of your code,
# not just the exported API functions you mention in
# workflow plan commands.

# Note: if you use `digest::digest()` instead of just `digest()`,
# `drake` does not dive into the function body anymore.
g <- function(x) {
  digest::digest(x) # Was previously just digest()
}
config <- drake_config(plan, cache = cache, history = FALSE)
tracked(config)
})

## End(Not run)

```

failed

List failed targets. **Maturing**

Description

Together, functions `failed()` and `diagnose()` should eliminate the strict need for ordinary error messages printed to the console.

Usage

```
failed(path = NULL, search = NULL, cache = drake::drake_cache(path =
  path, verbose = verbose), verbose = 1L, upstream_only = NULL)
```

Arguments

path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
search	Deprecated.
cache	drake cache. See <code>new_cache()</code> . If supplied, path is ignored.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
upstream_only	Deprecated.

Value

A character vector of target names.

See Also

`running()`, `make()`

Examples

```
## Not run:
isolate_example("contain side effects", {
  if (suppressWarnings(require("knitr"))){
    # Build a plan doomed to fail:
    bad_plan <- drake_plan(x = function_doesnt_exist())
    cache <- storrr::storrr_environment() # optional
    try(
      make(bad_plan, cache = cache, history = FALSE),
      silent = TRUE
    ) # error
    failed(cache = cache) # "x"
    e <- diagnose(x, cache = cache) # Retrieve the cached error log of x.
    names(e)
    e$error
    names(e$error)
  }
})

## End(Not run)
```

file_in

Declare input files and directories. **Stable**

Description

file_in() marks individual files (and whole directories) that your targets depend on.

Usage

```
file_in(...)
```

Arguments

... Character vector, paths to files and directories.

Value

A character vector of declared input file or directory paths.

URLs

As of drake 7.4.0, file_in() and file_out() have support for URLs. If the file name begins with "http://", "https://", or "ftp://", make() attempts to check the ETag to see if the data changed from last time. If no ETag can be found, drake simply uses the ETag from last make() and registers the file as unchanged (which prevents your workflow from breaking if you lose internet access). If your file_in() URLs require authentication, see the curl_handles argument of make() and drake_config() to learn how to supply credentials.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`file_out()`, `knitr_in()`, `ignore()`, `no_deps()`

Examples

```
## Not run:
isolate_example("contain side effects", {
  # The `file_out()` and `file_in()` functions
  # just takes in strings and returns them.
  file_out("summaries.txt")
  # Their main purpose is to orchestrate your custom files
  # in your workflow plan data frame.
  plan <- drake_plan(
    out = write.csv(mtcars, file_out("mtcars.csv")),
    contents = read.csv(file_in("mtcars.csv"))
  )
  plan
  # drake knows "\"mtcars.csv\"" is the first target
  # and a dependency of `contents`. See for yourself:

  make(plan)
  file.exists("mtcars.csv")

  # You can also work with entire directories this way.
  # However, in `file_out("your_directory")`, the directory
  # becomes an entire unit. Thus, `file_in("your_directory")`
  # is more appropriate for subsequent steps than
```

```
# `file_in("your_directory/file_inside.txt")`.
plan <- drake_plan(
  out = {
    dir.create(file_out("dir"))
    write.csv(mtcars, "dir/mtcars.csv")
  },
  contents = read.csv(file.path(file_in("dir"), "mtcars.csv"))
)
plan

make(plan)
file.exists("dir/mtcars.csv")

# See the connections that the file relationships create:
config <- drake_config(plan)
if (requireNamespace("visNetwork", quietly = TRUE)) {
  vis_drake_graph(config)
}
})

## End(Not run)
```

file_out

*Declare output files and directories. **Stable***

Description

file_out() marks individual files (and whole directories) that your targets create.

Usage

```
file_out(...)
```

Arguments

... Character vector, paths to files and directories.

Value

A character vector of declared output file or directory paths.

Keywords

drake_plan() understands special keyword functions for your commands. With the exception of target(), each one is a proper function with its own help file.

- **target()**: give the target more than just a command. Using **target()**, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), # nolint supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), # nolint or set any number of custom columns.
- **map()**: create multiple targets. Called inside **target()**.
- **split()**: create a target for each subset of data. Called inside **target()**.
- **cross()**: create a target for each combination of values. Called inside **target()**.

- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`file_out()`, `knitr_in()`, `ignore()`, `no_deps()`

Examples

```
## Not run:
isolate_example("contain side effects", {
# The `file_out()` and `file_in()` functions
# just takes in strings and returns them.
file_out("summaries.txt")
# Their main purpose is to orchestrate your custom files
# in your workflow plan data frame.
plan <- drake_plan(
  out = write.csv(mtcars, file_out("mtcars.csv")),
  contents = read.csv(file_in("mtcars.csv"))
)
plan
# drake knows "\"mtcars.csv\"" is the first target
# and a dependency of `contents`. See for yourself:

make(plan)
file.exists("mtcars.csv")

# You can also work with entire directories this way.
# However, in `file_out("your_directory")`, the directory
# becomes an entire unit. Thus, `file_in("your_directory")`
# is more appropriate for subsequent steps than
# `file_in("your_directory/file_inside.txt")`.
plan <- drake_plan(
  out = {
    dir.create(file_out("dir"))
    write.csv(mtcars, "dir/mtcars.csv")
  },
  contents = read.csv(file.path(file_in("dir"), "mtcars.csv"))
)
plan

make(plan)
file.exists("dir/mtcars.csv")

# See the connections that the file relationships create:
```

```

config <- drake_config(plan)
if (requireNamespace("visNetwork", quietly = TRUE)) {
  vis_drake_graph(config)
}
})

## End(Not run)

```

file_store

Show a file's encoded representation in the cache **Stable**

Description

This function simply wraps literal double quotes around the argument `x` so drake knows it is the name of a file. Use when you are calling functions like `deps_code()`: for example, `deps_code(file_store("report.m"))`. See the examples for details. Internally, drake wraps the names of file targets/imports inside literal double quotes to avoid confusion between files and generic R objects.

Usage

```
file_store(x)
```

Arguments

`x` Character string to be turned into a filename understandable by drake (i.e., a string with literal single quotes on both ends).

Value

A single-quoted character string: i.e., a filename understandable by drake.

Examples

```

# Wraps the string in single quotes.
file_store("my_file.rds") # "'my_file.rds'"
## Not run:
isolate_example("contain side effects", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the workflow to build the targets
    list.files() # Should include input "report.Rmd" and output "report.md".
    head(readadd(small)) # You can use symbols for ordinary objects.
    # But if you want to read cached info on files, use `file_store()`.
    readadd(file_store("report.md"), character_only = TRUE) # File fingerprint.
    deps_code(file_store("report.Rmd"))
    config <- drake_config(my_plan)
    deps_profile(
      file_store("report.Rmd"),
      config = config,
      character_only = TRUE
    )
  }
})

## End(Not run)

```

find_cache	<i>Search up the file system for the nearest drake cache.</i> Stable
------------	---

Description

Only works if the cache is a file system in a hidden folder named `.drake/` (default).

Usage

```
find_cache(path = getwd(), dir = NULL, directory = NULL)
```

Arguments

path	Starting path for search back for the cache. Should be a subdirectory of the drake project.
dir	Character, name of the folder containing the cache.
directory	Deprecated. Use dir.

Value

File path of the nearest drake cache or NULL if no cache is found.

See Also

[drake_plan\(\)](#), [make\(\)](#),

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the target.
    # Find the file path of the project's cache.
    # Search up through parent directories if necessary.
    find_cache()
  }
})

## End(Not run)
```

ignore	<i>Ignore code</i> Stable
--------	----------------------------------

Description

Ignore sections of commands and imported functions.

Usage

```
ignore(x = NULL)
```

Arguments

x Code to ignore.

Details

In user-defined functions and `drake_plan()` commands, you can wrap code chunks in `ignore()` to

1. Tell drake to not search for dependencies (targets etc. mentioned in the code) and
2. Ignore changes to the code so downstream targets remain up to date. To enforce (1) without (2), use `no_deps()`.

Value

The argument.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`file_in()`, `file_out()`, `knitr_in()`, `no_deps()`

Examples

```
## Not run:
isolate_example("Contain side effects", {
  # Normally, `drake` reacts to changes in dependencies.
  x <- 4
  make(plan = drake_plan(y = sqrt(x)))
  x <- 5
  make(plan = drake_plan(y = sqrt(x)))
  make(plan = drake_plan(y = sqrt(4) + x))
  # But not with ignore().
  make(plan = drake_plan(y = sqrt(4) + ignore(x))) # Builds y.
  x <- 6
  make(plan = drake_plan(y = sqrt(4) + ignore(x))) # Skips y.
  make(plan = drake_plan(y = sqrt(4) + ignore(x + 1))) # Skips y.

  # ignore() works with functions and multiline code chunks.
  f <- function(x) {
    ignore({
      x <- x + 1
      x <- x + 2
    })
    x # Not ignored.
  }
  make(plan = drake_plan(y = f(2)))
  readd(x)
  # Changes the content of the ignore() block:
  f <- function(x) {
    ignore({
      x <- x + 1
    })
    x # Not ignored.
  }
  make(plan = drake_plan(x = f(2)))
  readd(x)
})

## End(Not run)
```

knitr_in

Declare knitr/rmarkdown source files as dependencies. **Stable**

Description

knitr_in() marks individual knitr/R Markdown reports as dependencies. In drake, these reports are pieces of the pipeline. R Markdown is a great tool for *displaying* precomputed results, but not for running a large workflow from end to end. These reports should do as little computation as possible.

Usage

```
knitr_in(...)
```

Arguments

... Character strings. File paths of knitr/rmarkdown source files supplied to a command in your workflow plan data frame.

Details

Unlike `file_in()` and `file_out()`, `knitr_in()` does not work with entire directories.

Value

A character vector of declared input file paths.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`file_in()`, `file_out()`, `ignore()`, `no_deps()`

Examples

```
## Not run:
isolate_example("contain side effects", {
  # `knitr_in()` is like `file_in()`
  # except that it analyzes active code chunks in your `knitr`
  # source file and detects non-file dependencies.
  # That way, updates to the right dependencies trigger rebuilds
  # in your report.
  # The mtcars example (`drake_example("mtcars")`)
  # already has a demonstration
```

```

load_mtcars_example()
make(my_plan)

# Now how did drake magically know that
# `small`, `large`, and `coef_regression2_small` were
# dependencies of the output file `report.md`?
# because the command in the workflow plan had
# `knitr_in("report.Rmd")` in it, so drake knew
# to analyze the active code chunks. There, it spotted
# where `small`, `large`, and `coef_regression2_small`
# were read from the cache using calls to `loadadd()` and `readd()`.
})

## End(Not run)

```

legend_nodes	Create the nodes data frame used in the legend of the graph visualizations. Soft-deprecated
--------------	--

Description

Output a visNetwork-friendly data frame of nodes. It tells you what the colors and shapes mean in the graph visualizations.

Usage

```
legend_nodes(font_size = 20)
```

Arguments

font_size Font size of the node label text.

Value

A data frame of legend nodes for the graph visualizations.

Examples

```

## Not run:
# Show the legend nodes used in graph visualizations.
# For example, you may want to inspect the color palette more closely.
if (requireNamespace("visNetwork", quietly = TRUE)) {
  # visNetwork::visNetwork(nodes = legend_nodes()) # nolint
}

## End(Not run)

```

load_mtcars_example *Load the mtcars example. **Maturing***

Description

Is there an association between the weight and the fuel efficiency of cars? To find out, we use the mtcars example from `drake_example("mtcars")`. The mtcars dataset itself only has 32 rows, so we generate two larger bootstrapped datasets and then analyze them with regression models. Finally, we summarize the regression models to see if there is an association.

Usage

```
load_mtcars_example(envir = parent.frame(), report_file = NULL,
  overwrite = FALSE, force = FALSE)
```

Arguments

<code>envir</code>	The environment to load the example into. Defaults to your workspace. For an insulated workspace, set <code>envir = new.env(parent = globalenv())</code> .
<code>report_file</code>	Where to write the report file. Deprecated. In a future release, the report file will always be <code>report.Rmd</code> and will always be written to your working directory (current default).
<code>overwrite</code>	Logical, whether to overwrite an existing file <code>report.Rmd</code> .
<code>force</code>	Deprecated.

Details

Use `drake_example("mtcars")` to get the code for the mtcars example. This function also writes/overwrites the file, `report.Rmd`.

Value

Nothing.

See Also

[clean_mtcars_example\(\)](#) [drake_examples\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    # Populate your workspace and write 'report.Rmd'.
    load_mtcars_example() # Get the code: drake_example("mtcars")
    # Check the dependencies of an imported function.
    deps_code(reg1)
    # Check the dependencies of commands in the workflow plan.
    deps_code(my_plan$command[1])
    deps_code(my_plan$command[4])
    # Plot the interactive network visualization of the workflow.
    config <- drake_config(my_plan)
```



```

outdated(config) # Which targets are out of date?
# Run the workflow to build all the targets in the plan.
make(my_plan)
outdated(config) # Everything should be up to date.
# For the reg2() model on the small dataset,
# the p-value is so small that there may be an association
# between weight and fuel efficiency after all.
readd(coef_regression2_small)
# Clean up the example.
clean_mtcars_example()
}
})

## End(Not run)

```

make

Run your project (build the outdated targets). Maturing

Description

This is the central, most important function of the drake package. It runs all the steps of your workflow in the correct order, skipping any work that is already up to date. See <https://github.com/ropensci/drake/blob/master/README.md#documentation> for an overview of the documentation.

Usage

```

make(plan, targets = NULL, envir = parent.frame(), verbose = 1L,
  hook = NULL, cache = drake::drake_cache(verbose = verbose,
    console_log_file = console_log_file), fetch_cache = NULL,
  parallelism = "loop", jobs = 1L, jobs_preprocess = 1L,
  packages = rev(.packages()), lib_loc = NULL,
  prework = character(0), prepend = NULL, command = NULL,
  args = NULL, recipe_command = NULL, log_progress = TRUE,
  skip_targets = FALSE, timeout = NULL, cpu = Inf, elapsed = Inf,
  retries = 0, force = FALSE, graph = NULL,
  trigger = drake::trigger(), skip_imports = FALSE,
  skip_safety_checks = FALSE, config = NULL, lazy_load = "eager",
  session_info = TRUE, cache_log_file = NULL, seed = NULL,
  caching = "master", keep_going = FALSE, session = NULL,
  pruning_strategy = NULL, makefile_path = NULL,
  console_log_file = NULL, ensure_workers = NULL,
  garbage_collection = FALSE, template = list(), sleep = function(i)
  0.01, hasty_build = NULL, memory_strategy = "speed", layout = NULL,
  lock_envir = TRUE, history = TRUE, recover = FALSE,
  recoverable = TRUE, curl_handles = list())

```

Arguments

plan	Workflow plan data frame. A workflow plan data frame is a data frame with a target column and a command column. (See the details in the drake_plan() help file for descriptions of the optional columns.) Targets are the objects that
------	--

	drake generates, and commands are the pieces of R code that produce them. You can create and track custom files along the way (see <code>file_in()</code> , <code>file_out()</code> , and <code>knitr_in()</code>). Use the function <code>drake_plan()</code> to generate workflow plan data frames.
targets	Character vector, names of targets to build. Dependencies are built too. Together, the plan and targets comprise the workflow network (i.e. the graph argument). Changing either will change the network.
envir	Environment to use. Defaults to the current workspace, so you should not need to worry about this most of the time. A deep copy of <code>envir</code> is made, so you don't need to worry about your workspace being modified by <code>make</code> . The deep copy inherits from the global environment. Wherever necessary, objects and functions are imported from <code>envir</code> and the global environment and then reproducibly tracked as dependencies.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
hook	Deprecated.
cache	drake cache as created by <code>new_cache()</code> . See also <code>drake_cache()</code> .
fetch_cache	Deprecated.
parallelism	Character scalar, type of parallelism to use. For detailed explanations, see the high-performance computing chapter # nolint of the user manual. You could also supply your own scheduler function if you want to experiment or aggressively optimize. The function should take a single config argument (produced by <code>drake_config()</code>). Existing examples from drake's internals are the <code>backend_*()</code> functions: <ul style="list-style-type: none"> • <code>backend_loop()</code> • <code>backend_clustermq()</code> • <code>backend_future()</code> However, this functionality is really a back door and should not be used for production purposes unless you really know what you are doing and you are willing to suffer setbacks whenever drake's un-exported core functions are updated.
jobs	Maximum number of parallel workers for processing the targets. You can experiment with <code>predict_runtime()</code> to help decide on an appropriate number of jobs. For details, visit https://ropenscilabs.github.io/drake-manual/time.html .
jobs_preprocess	Number of parallel jobs for processing the imports and doing other preprocessing tasks.
packages	Character vector packages to load, in the order they should be loaded. Defaults to <code>rev(.packages())</code> , so you should not usually need to set this manually. Just call <code>library()</code> to load your packages before <code>make()</code> . However, sometimes packages need to be strictly forced to load in a certain order, especially if <code>parallelism</code> is "Makefile". To do this, do not use <code>library()</code> or <code>require()</code> or <code>loadNamespace()</code> or <code>attachNamespace()</code> to load any libraries beforehand. Just list your packages in the <code>packages</code> argument in the order you want them to be loaded.
lib_loc	Character vector, optional. Same as in <code>library()</code> or <code>require()</code> . Applies to the <code>packages</code> argument (see above).

prework	Expression (language object), list of expressions, or character vector. Code to run right before targets build. Called only once if <code>parallelism</code> is "loop" and once per target otherwise. This code can be used to set global options, etc.
prepend	Deprecated.
command	Deprecated.
args	Deprecated.
recipe_command	Deprecated.
log_progress	Logical, whether to log the progress of individual targets as they are being built. Progress logging creates extra files in the cache (usually the <code>.drake/</code> folder) and slows down <code>make()</code> a little. If you need to reduce or limit the number of files in the cache, call <code>make(log_progress = FALSE, recover = FALSE)</code> .
skip_targets	Logical, whether to skip building the targets in plan and just import objects and files.
timeout	deprecated. Use <code>elapsed</code> and <code>cpu</code> instead.
cpu	Same as the <code>cpu</code> argument of <code>setTimeLimit()</code> . Seconds of cpu time before a target times out. Assign target-level cpu timeout times with an optional <code>cpu</code> column in plan.
elapsed	Same as the <code>elapsed</code> argument of <code>setTimeLimit()</code> . Seconds of elapsed time before a target times out. Assign target-level elapsed timeout times with an optional <code>elapsed</code> column in plan.
retries	Number of retries to execute if the target fails. Assign target-level retries with an optional <code>retries</code> column in plan.
force	Logical. If <code>FALSE</code> (default) then drake imposes checks if the cache was created with an old and incompatible version of drake. If there is an incompatibility, <code>make()</code> stops to give you an opportunity to downgrade drake to a compatible version rather than rerun all your targets from scratch.
graph	An <code>igraph</code> object from the previous <code>make()</code> . Supplying a pre-built graph could save time.
trigger	Name of the trigger to apply to all targets. Ignored if plan has a <code>trigger</code> column. See trigger() for details.
skip_imports	Logical, whether to totally neglect to process the imports and jump straight to the targets. This can be useful if your imports are massive and you just want to test your project, but it is bad practice for reproducible data analysis. This argument is overridden if you supply your own <code>graph</code> argument.
skip_safety_checks	Logical, whether to skip the safety checks on your workflow. Use at your own peril.
config	A list generated by drake_config() . If you supply <code>config</code> to <code>make()</code> , you must not supply any additional arguments. drake_config() collects and sanitizes the multitude of parameters and settings that <code>make()</code> needs to do its job: the plan, packages, the environment of functions and initial data objects, parallel computing instructions, verbosity level, etc. Once you generate a list with drake_config() , do not modify it by hand. Utility functions such as outdated() , vis_drake_graph() , and predict_runtime() require output from drake_config() for the <code>config</code> argument.
lazy_load	Either a character vector or a logical. Choices: <ul style="list-style-type: none"> • "eager": no lazy loading. The target is loaded right away with assign().

- "promise": lazy loading with `delayedAssign()`
- "bind": lazy loading with active bindings: `bindr::populate_env()`.
- TRUE: same as "promise".
- FALSE: same as "eager".

`lazy_load` should not be "promise" for "parLapply" parallelism combined with jobs greater than 1. For local multi-session parallelism and lazy loading, try `library(future); future::plan(multisession)` and then `make(..., parallelism = "future_lapply", lazy_load = "bind")`.

If `lazy_load` is "eager", drake prunes the execution environment before each target/stage, removing all superfluous targets and then loading any dependencies it will need for building. In other words, drake prepares the environment in advance and tries to be memory efficient. If `lazy_load` is "bind" or "promise", drake assigns promises to load any dependencies at the last minute. Lazy loading may be more memory efficient in some use cases, but it may duplicate the loading of dependencies, costing time.

session_info Logical, whether to save the `sessionInfo()` to the cache. This behavior is recommended for serious `make()`s for the sake of reproducibility. This argument only exists to speed up tests. Apparently, `sessionInfo()` is a bottleneck for small `make()`s.

cache_log_file Name of the CSV cache log file to write. If TRUE, the default file name is used (`drake_cache.CSV`). If NULL, no file is written. If activated, this option writes a flat text file to represent the state of the cache (fingerprints of all the targets and imports). If you put the log file under version control, your commit history will give you an easy representation of how your results change over time as the rest of your project changes. Hopefully, this is a step in the right direction for data reproducibility.

seed Integer, the root pseudo-random number generator seed to use for your project. In `make()`, drake generates a unique local seed for each target using the global seed and the target name. That way, different pseudo-random numbers are generated for different targets, and this pseudo-randomness is reproducible.

To ensure reproducibility across different R sessions, `set.seed()` and `.Random.seed` are ignored and have no affect on drake workflows. Conversely, `make()` does not usually change `.Random.seed`, even when pseudo-random numbers are generated. The exception to this last point is `make(parallelism = "clustermq")` because the `clustermq` package needs to generate random numbers to set up ports and sockets for ZeroMQ.

On the first call to `make()` or `drake_config()`, drake uses the random number generator seed from the `seed` argument. Here, if the seed is NULL (default), drake uses a seed of 0. On subsequent `make()`s for existing projects, the project's cached seed will be used in order to ensure reproducibility. Thus, the `seed` argument must either be NULL or the same seed from the project's cache (usually the `.drake/` folder). To reset the random number generator seed for a project, use `clean(destroy = TRUE)`.

caching Character string, either "master" or "worker".

- "master": Targets are built by remote workers and sent back to the master process. Then, the master process saves them to the cache (`config$cache`, usually a file system `storr`). Appropriate if remote workers do not have access to the file system of the calling R session. Targets are cached one at a time, which may be slow in some situations.

	<ul style="list-style-type: none"> • "worker": Remote workers not only build the targets, but also save them to the cache. Here, caching happens in parallel. However, remote workers need to have access to the file system of the calling R session. Transferring target data across a network can be slow.
keep_going	Logical, whether to still keep running <code>make()</code> if targets fail.
session	Deprecated. Has no effect now.
pruning_strategy	Deprecated. See <code>memory_strategy</code> .
makefile_path	Deprecated.
console_log_file	Optional character scalar of a file name or connection object (such as <code>stdout()</code>) to dump maximally verbose log information for <code>make()</code> and other functions (all functions that accept a <code>config</code> argument, plus <code>drake_config()</code>). If you choose to use a text file as the console log, it will persist over multiple function calls until you delete it manually. Fields in each row the log file, from left to right: - The node name (short host name) of the computer (from <code>Sys.info()["nodename"]</code>). - The process ID (from <code>Sys.getpid()</code>). - A timestamp with the date and time (in microseconds). - A brief description of what drake was doing. The fields are separated by pipe symbols (" ").
ensure_workers	Deprecated.
garbage_collection	Logical, whether to call <code>gc()</code> each time a target is built during <code>make()</code> .
template	A named list of values to fill in the <code>{{ ... }}</code> placeholders in template files (e.g. from <code>drake_hpc_template_file()</code>). Same as the <code>template</code> argument of <code>clustermq::Q()</code> and <code>clustermq::workers</code> . Enabled for <code>clustermq</code> only (<code>make(parallelism = "clustermq")</code>), not <code>future</code> or <code>batchtools</code> so far. For more information, see the <code>clustermq</code> package: https://github.com/mschubert/clustermq . Some template placeholders such as <code>{{ job_name }}</code> and <code>{{ n_jobs }}</code> cannot be set this way.
sleep	<p>Optional function on a single numeric argument <code>i</code>. Default: <code>function(i) 0.01</code>. To conserve memory, drake assigns a brand new closure to <code>sleep</code>, so your custom function should not depend on in-memory data except from loaded packages.</p> <p>For parallel processing, drake uses a central master process to check what the parallel workers are doing, and for the affected high-performance computing workflows, wait for data to arrive over a network. In between loop iterations, the master process sleeps to avoid throttling. The <code>sleep</code> argument to <code>make()</code> and <code>drake_config()</code> allows you to customize how much time the master process spends sleeping.</p> <p>The <code>sleep</code> argument is a function that takes an argument <code>i</code> and returns a numeric scalar, the number of seconds to supply to <code>Sys.sleep()</code> after iteration <code>i</code> of checking. (Here, <code>i</code> starts at 1.) If the checking loop does something other than sleeping on iteration <code>i</code>, then <code>i</code> is reset back to 1.</p> <p>To sleep for the same amount of time between checks, you might supply something like <code>function(i) 0.01</code>. But to avoid consuming too many resources during heavier and longer workflows, you might use an exponential back-off: say, <code>function(i) { 0.1 + 120 * pexp(i - 1, rate = 0.01) }</code>.</p>
hasty_build	A user-defined function. In "hasty mode" (<code>make(parallelism = "hasty")</code>) this is the function that evaluates a target's command and returns the resulting value.

The `hasty_build` argument has no effect if `parallelism` is any value other than "hasty".

The function you pass to `hasty_build` must have arguments `target` and `config`. Here, `target` is a character scalar naming the target being built, and `config` is a configuration list of runtime parameters generated by `drake_config()`.

`memory_strategy`

Character scalar, name of the strategy drake uses to load/unload a target's dependencies in memory. You can give each target its own memory strategy, (e.g. `drake_plan(x = 1, y = target(f(x), memory_strategy = "lookahead"))`) to override the global memory strategy. Choices:

- "speed": Once a target is newly built or loaded in memory, just keep it there. This choice maximizes speed and hogs memory.
- "autoclean": Just before building each new target, unload everything from memory except the target's direct dependencies. After a target is built, discard it from memory. (Set `garbage_collection = TRUE` to make sure it is really gone.) This option conserves memory, but it sacrifices speed because each new target needs to reload any previously unloaded targets from storage.
- "preclean": Just before building each new target, unload everything from memory except the target's direct dependencies. After a target is built, keep it in memory until drake determines they can be unloaded. This option conserves memory, but it sacrifices speed because each new target needs to reload any previously unloaded targets from storage.
- "lookahead": Just before building each new target, search the dependency graph to find targets that will not be needed for the rest of the current `make()` session. After a target is built, keep it in memory until the next memory management stage. In this mode, targets are only in memory if they need to be loaded, and we avoid superfluous reads from the cache. However, searching the graph takes time, and it could even double the computational overhead for large projects.
- "unload": Just before building each new target, unload all targets from memory. After a target is built, **do not** keep it in memory. This mode aggressively optimizes for both memory and speed, but in commands and triggers, you have to manually load any dependencies you need using `readd()`.
- "none": Do not manage memory at all. Do not load or unload anything before building targets. After a target is built, **do not** keep it in memory. This mode aggressively optimizes for both memory and speed, but in commands and triggers, you have to manually load any dependencies you need using `readd()`.

For even more direct control over which targets drake keeps in memory, see the help file examples of `drake_envir()`. Also see the `garbage_collection` argument of `make()` and `drake_config()`.

`layout`

`config$layout`, where `config` is the return value from a prior call to `drake_config()`. If your plan or environment have changed since the last `make()`, do not supply a `layout` argument. Otherwise, supplying one could save time.

`lock_envir`

Logical, whether to lock `config$envir` during `make()`. If `TRUE`, `make()` quits in error whenever a command in your drake plan (or prework) tries to add, remove, or modify non-hidden variables in your environment/workspace/R session. This is extremely important for ensuring the purity of your functions and the reproducibility/credibility/trust you can place in your project. `lock_envir` will be set to a default of `TRUE` in drake version 7.0.0 and higher.

history	Logical, whether to record the build history of your targets. You can also supply a <code>txtq</code> , which is how drake records history. Must be TRUE for <code>drake_history()</code> to work later.
recover	<p>Logical, whether to activate automated data recovery. The default is FALSE because</p> <ol style="list-style-type: none"> 1. Automated data recovery is still experimental. 2. It has reproducibility issues. Targets recovered from the distant past may have been generated with earlier versions of R and earlier package environments that no longer exist. <p>How it works: if <code>recover</code> is TRUE, drake tries to salvage old target values from the cache instead of running commands from the plan. A target is recoverable if</p> <ol style="list-style-type: none"> 1. There is an old value somewhere in the cache that shares the command, dependencies, etc. of the target about to be built. 2. The old value was generated with <code>make(recoverable = TRUE)</code>. <p>If both conditions are met, drake will</p> <ol style="list-style-type: none"> 1. Assign the most recently-generated admissible data to the target, and 2. skip the target's command. <p>Functions <code>recoverable()</code> and <code>r_recoverable()</code> show the most upstream outdated targets that will be recovered in this way in the next <code>make()</code> or <code>r_make()</code>.</p>
recoverable	Logical, whether to make target values recoverable with <code>make(recover = TRUE)</code> . This requires writing extra files to the cache, and it prevents old metadata from being removed with garbage collection (<code>clean(garbage_collection = TRUE)</code> , <code>gc()</code> in <code>storrs</code>). If you need to limit the cache size or the number of files in the cache, consider <code>make(recoverable = FALSE, progress = FALSE)</code> .
curl_handles	<p>A named list of curl handles. Each value is an object from <code>curl::new_handle()</code>, and each name is a URL (and should start with "http", "https", or "ftp"). Example: <code>list(http://httpbin.org/basic-auth = curl::new_handle(username = "user", password = "passwd"))</code> Then, if your plan has <code>file_in("http://httpbin.org/basic-auth/...")</code> drake will authenticate using the username and password of the handle for <code>http://httpbin.org/basic-auth/</code>.</p> <p>drake uses partial matching on text to find the right handle of the <code>file_in()</code> URL, so the name of the handle could be the complete URL ("<code>http://httpbin.org/basic-auth/user</code>") or a part of the URL (e.g. "<code>http://httpbin.org/</code>" or "<code>http://httpbin.org/basic-auth/</code>"). If you have multiple handles whose names match your URL, drake will choose the closest match.</p>

Value

nothing

Interactive mode

In interactive sessions, consider `r_make()`, `r_outdated()`, etc. rather than `make()`, `outdated()`, etc. The `r_*`() drake functions are more reproducible when the session is interactive.

A serious drake workflow should be consistent and reliable, ideally with the help of a master R script. This script should begin in a fresh R session, load your packages and functions in a dependable manner, and then run `make()`. Example: <https://github.com/wlandau/drake-examples/tree/master/gsp>. Batch mode, especially within a container, is particularly helpful.

Interactive R sessions are still useful, but they easily grow stale. Targets can falsely invalidate if you accidentally change a function or data object in your environment.

Self-invalidation

It is possible to construct a workflow that tries to invalidate itself. Example:

```
plan <- drake_plan(
  x = {
    data(mtcars)
    mtcars$mpg
  },
  y = mean(x)
)
```

Here, because `data()` loads `mtcars` into the global environment, the very act of building `x` changes the dependencies of `x`. In other words, without safeguards, `x` would not be up to date at the end of `make(plan)`. Please try to avoid workflows that modify the global environment. Functions such as `data()` belong in your setup scripts prior to `make()`, not in any functions or commands that get called during `make()` itself.

For each target that is still problematic (e.g. <https://github.com/rstudio/gt/issues/297>) you can safely run the command in its own special `callr::r()` process. Example: <https://github.com/rstudio/gt/issues/297#issuecomment-497778735>. # nolint

If that fails, you can run `make(plan, lock_envir = FALSE)` to suppress environment-locking for all targets. However, this is not usually recommended. There are legitimate use cases for `lock_envir = FALSE` (example: <https://ropenscilabs.github.io/drake-manual/hpc.html#parallel-computing-within-targets>) # nolint but most workflows should stick with the default `lock_envir = TRUE`.

See Also

[drake_plan\(\)](#), [drake_config\(\)](#), [vis_drake_graph\(\)](#), [outdated\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    config <- drake_config(my_plan)
    outdated(config) # Which targets need to be (re)built?
    make(my_plan) # Build what needs to be built.
    outdated(config) # Everything is up to date.
    # Change one of your imported function dependencies.
    reg2 = function(d) {
      d$x3 = d$x^3
      lm(y ~ x3, data = d)
    }
    outdated(config) # Some targets depend on reg2().
    make(my_plan) # Rebuild just the outdated targets.
    outdated(config) # Everything is up to date again.
    if (requireNamespace("visNetwork", quietly = TRUE)) {
      vis_drake_graph(config) # See how they fit in an interactive graph.
      make(my_plan, cache_log_file = TRUE) # Write a CSV log file this time.
      vis_drake_graph(config) # The colors changed in the graph.
      # Run targets in parallel:
      # options(clustermq.scheduler = "multicore") # nolint
      # make(my_plan, parallelism = "clustermq", jobs = 2) # nolint
    }
  }
})
```



```

clean() # Start from scratch next time around.
}
})

## End(Not run)

```

map

Define multiple targets at once **Maturing****Description**

Similar to `pmap()` from `purrr`, except `drake`'s `map()` defines new targets.

Usage

```
map(..., .data, .id, .tag_in, .tag_out)
```

Arguments

<code>...</code>	Grouping variables. New grouping variables must be supplied with their names and values, existing grouping variables can be given as symbols without any values assigned.
<code>.data</code>	A data frame of new grouping variables with grouping variable names as column names and values as elements.
<code>.id</code>	Symbol or vector of symbols naming grouping variables to incorporate into target names. Useful for creating short target names. Set <code>.id = FALSE</code> to use integer indices as target name suffixes.
<code>.tag_in</code>	A symbol or vector of symbols. Tags assign targets to grouping variables. Use <code>.tag_in</code> to assign <i>untransformed</i> targets to grouping variables.
<code>.tag_out</code>	Just like <code>.tag_in</code> , except that <code>.tag_out</code> assigns <i>transformed</i> targets to grouping variables.

Details

Only valid within a call to `target()` in `drake_plan()`. See the examples below.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.

- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

split, cross, combine, drake_plan, target

Examples

```
models <- c("glm", "hierarchical")
plan <- drake_plan(
  data = target(
    get_data(x),
    transform = map(x = c("simulated", "survey"))
  ),
  analysis = target(
    analyze_data(data, model),
    transform = cross(data, model = !!models, .id = c(x, model))
  ),
  summary = target(
    summarize_analysis(analysis),
    transform = map(analysis, .id = c(x, model))
  ),
  results = target(
    bind_rows(summary),
    transform = combine(summary, .by = data)
  )
)
plan
if (requireNamespace("styler")) {
  print(drake_plan_source(plan))
}
# Tags:
drake_plan(
  x = target(
    command,
    transform = map(y = c(1, 2), .tag_in = from, .tag_out = c(to, out))
  ),
  trace = TRUE
)
plan <- drake_plan(
  survey = target(
    survey_data(x),
    transform = map(x = c(1, 2), .tag_in = source, .tag_out = dataset)
  ),
  download = target(
    download_data(),
```

```

    transform = map(y = c(5, 6), .tag_in = source, .tag_out = dataset)
  ),
  analysis = target(
    analyze(dataset),
    transform = map(dataset)
  ),
  results = target(
    bind_rows(analysis),
    transform = combine(analysis, .by = source)
  )
)
plan
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}

```

missed

*Report any import objects required by your drake_plan plan but missing from your workspace or file system. **Stable***

Description

Checks your workspace/environment and file system.

Usage

```
missed(config)
```

Arguments

config Internal runtime parameter list produced by both `drake_config()` and `make()`.

Value

Character vector of names of missing objects and files.

See Also

[outdated\(\)](#)

Examples

```

## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    config <- drake_config(my_plan)
    missed(config) # All the imported files and objects should be present.
    rm(reg1) # Remove an import dependency from you workspace.
    missed(config) # Should report that reg1 is missing.
  }
})

## End(Not run)

```

new_cache

Make a new drake cache. **Maturing****Description**

Uses the [storr_rds\(\)](#) function from the storr package.

Usage

```
new_cache(path = NULL, verbose = 1L, type = NULL,
          hash_algorithm = NULL, short_hash_algo = NULL,
          long_hash_algo = NULL, ..., console_log_file = NULL)
```

Arguments

path	File path to the cache if the cache is a file system cache.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
type	Deprecated argument. Once stood for cache type. Use storr to customize your caches instead.
hash_algorithm	Name of a hash algorithm to use. See the algo argument of the digest package for your options.
short_hash_algo	Deprecated on 2018-12-12. Use hash_algorithm instead.
long_hash_algo	Deprecated on 2018-12-12. Use hash_algorithm instead.
...	other arguments to the cache constructor.
console_log_file	Optional character scalar of a file name or connection object (such as stdout()) to dump maximally verbose log information for make() and other functions (all functions that accept a config argument, plus drake_config()). If you choose to use a text file as the console log, it will persist over multiple function calls until you delete it manually. Fields in each row the log file, from left to right: - The node name (short host name) of the computer (from Sys.info()["nodename"]). - The process ID (from Sys.getpid()). - A timestamp with the date and time (in microseconds). - A brief description of what drake was doing. The fields are separated by pipe symbols (" ").

Value

A newly created drake cache as a storr object.

See Also

[make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine new_cache() side effects.", {
  clean(destroy = TRUE) # Should not be necessary.
  unlink("not_hidden", recursive = TRUE) # Should not be necessary.
  cache1 <- new_cache() # Creates a new hidden '.drake' folder.
  cache2 <- new_cache(path = "not_hidden", hash_algorithm = "md5")
  clean(destroy = TRUE, cache = cache2)
})

## End(Not run)
```

no_deps

Suppress dependency detection. Stable

Description

Tell drake to not search for dependencies in a chunk of code.

Usage

```
no_deps(x = NULL)
```

Arguments

x Code for which dependency detection is suppressed.

Details

`no_deps()` is similar to `ignore()`, but it still lets drake track meaningful changes to the code itself.

Value

The argument.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), `# nolint` supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), `# nolint` or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.

- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`file_in()`, `file_out()`, `knitr_in()`, `no_deps()`

Examples

```
## Not run:
isolate_example("Contain side effects", {
  # Normally, `drake` reacts to changes in dependencies.
  x <- 4
  make(plan = drake_plan(y = sqrt(x)))
  x <- 5
  make(plan = drake_plan(y = sqrt(x)))
  make(plan = drake_plan(y = sqrt(4) + x))
  # But not with no_deps().
  make(plan = drake_plan(y = sqrt(4) + no_deps(x))) # Builds y.
  x <- 6
  make(plan = drake_plan(y = sqrt(4) + no_deps(x))) # Skips y.
  # However, `drake` *does* react to changes
  # to the *literal code* inside `no_deps()`.
  make(plan = drake_plan(y = sqrt(4) + ignore(x + 1))) # Builds y.

  # Like ignore(), no_deps() works with functions and multiline code chunks.
  z <- 1
  f <- function(x) {
    no_deps({
      x <- z + 1
      x <- x + 2
    })
    x
  }
  make(plan = drake_plan(y = f(2)))
  readd(y)
  z <- 2 # Changed dependency is not tracked.
  make(plan = drake_plan(y = f(2)))
  readd(y)
})

## End(Not run)
```

Description

Outdated targets will be rebuilt in the next `make()`.

Usage

```
outdated(config, make_imports = TRUE, do_prework = TRUE)
```

Arguments

<code>config</code>	Optional internal runtime parameter list produced with <code>drake_config()</code> . You must use a fresh <code>config</code> argument with an up-to-date dependency graph that was never modified by hand. If needed, rerun <code>drake_config()</code> early and often. See the details in the help file for <code>drake_config()</code> .
<code>make_imports</code>	Logical, whether to make the imports first. Set to <code>FALSE</code> to save some time and risk obsolete output.
<code>do_prework</code>	Whether to do the prework normally supplied to <code>make()</code> .

Value

Character vector of the names of outdated targets.

See Also

`r_outdated()`, `drake_config()`, `missed()`, `drake_plan()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    # Recompute the config list early and often to have the
    # most current information. Do not modify the config list by hand.
    config <- drake_config(my_plan)
    outdated(config = config) # Which targets are out of date?
    make(my_plan) # Run the projects, build the targets.
    config <- drake_config(my_plan)
    # Now, everything should be up to date (no targets listed).
    outdated(config = config)
  }
})

## End(Not run)
```

Description

`code_to_plan()`, `plan_to_code()`, and `plan_to_notebook()` together illustrate the relationships between drake plans, R scripts, and R Markdown documents. In the file generated by `plan_to_code()`, every target/command pair becomes a chunk of code. Targets are arranged in topological order so dependencies are available before their downstream targets. Please note:

1. You are still responsible for loading your project's packages, imported functions, etc.
2. Triggers disappear.

Usage

```
plan_to_code(plan, con = stdout())
```

Arguments

<code>plan</code>	Workflow plan data frame. See <code>drake_plan()</code> for details.
<code>con</code>	A file path or connection to write to.

See Also

`drake_plan()`, `make()`, `code_to_plan()`, `plan_to_notebook()`

Examples

```
plan <- drake_plan(
  raw_data = read_excel(file_in("raw_data.xlsx")),
  data = raw_data,
  hist = create_plot(data),
  fit = lm(Sepal.Width ~ Petal.Width + Species, data)
)
file <- tempfile()
# Turn the plan into an R script at the given file path.
plan_to_code(plan, file)
# Here is what the script looks like.
cat(readLines(file), sep = "\n")
# Convert back to a drake plan.
code_to_plan(file)
```

<code>plan_to_notebook</code>	<i>Turn a drake plan into an R notebook.</i> Stable
-------------------------------	--

Description

`code_to_plan()`, `plan_to_code()`, and `plan_to_notebook()` together illustrate the relationships between drake plans, R scripts, and R Markdown documents. In the file generated by `plan_to_code()`, every target/command pair becomes a chunk of code. Targets are arranged in topological order so dependencies are available before their downstream targets. Please note:

1. You are still responsible for loading your project's packages, imported functions, etc.
2. Triggers disappear.

Usage

```
plan_to_notebook(plan, con)
```

Arguments

`plan` Workflow plan data frame. See [drake_plan\(\)](#) for details.

`con` A file path or connection to write to.

See Also

[drake_plan\(\)](#), [make\(\)](#), [code_to_plan\(\)](#), [plan_to_code\(\)](#)

Examples

```
if (suppressWarnings(require("knitr"))){
  plan <- drake_plan(
    raw_data = read_excel(file_in("raw_data.xlsx")),
    data = raw_data,
    hist = create_plot(data),
    fit = lm(Sepal.Width ~ Petal.Width + Species, data)
  )
  file <- tempfile()
  # Turn the plan into an R notebook at the given file path.
  plan_to_notebook(plan, file)
  # Here is what the script looks like.
  cat(readLines(file), sep = "\n")
  # Convert back to a drake plan.
  code_to_plan(file)
}
```

predict_runtime	<i>Predict the elapsed runtime of the next call to <code>make()</code> for non-staged parallel backends. Maturing</i>
-----------------	--

Description

Take the past recorded runtimes times from [build_times\(\)](#) and use them to predict how the targets will be distributed among the available workers in the next [make\(\)](#). Then, predict the overall runtime to be the runtime of the slowest (busiest) workers. Predictions only include the time it takes to run the targets, not overhead/preprocessing from drake itself.

Usage

```
predict_runtime(config, targets = NULL, from_scratch = FALSE,
  targets_only = NULL, jobs = 1, known_times = numeric(0),
  default_time = 0, warn = TRUE)
```

Arguments

config	Optional internal runtime parameter list of produced by both <code>make()</code> and <code>drake_config()</code> .
targets	Character vector, names of targets. Predict the runtime of building these targets plus dependencies. Defaults to all targets.
from_scratch	Logical, whether to predict a <code>make()</code> build from scratch or to take into account the fact that some targets may be already up to date and therefore skipped.
targets_only	Deprecated.
jobs	The <code>jobs</code> argument of your next planned <code>make()</code> . How many targets to do you plan to have running simultaneously?
known_times	A named numeric vector with targets/imports as names and values as hypothetical runtimes in seconds. Use this argument to overwrite any of the existing build times or the <code>default_time</code> .
default_time	Number of seconds to assume for any target or import with no recorded runtime (from <code>build_times()</code>) or anything in <code>known_times</code> .
warn	Logical, whether to warn the user about any targets with no available runtime, either in <code>known_times</code> or <code>build_times()</code> . The times for these targets default to <code>default_time</code> .

Value

Predicted total runtime of the next call to `make()`.

See Also

`predict_workers()`, `build_times()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    config <- drake_config(my_plan)
    known_times <- rep(7200, nrow(my_plan))
    names(known_times) <- my_plan$target
    known_times
    # Predict the runtime
    if (requireNamespace("lubridate", quietly = TRUE)) {
      predict_runtime(
        config,
        jobs = 7,
        from_scratch = TRUE,
        known_times = known_times
      )
      predict_runtime(
        config,
        jobs = 8,
        from_scratch = TRUE,
        known_times = known_times
      )
    }
    balance <- predict_workers(
      config,
```

```

    jobs = 7,
    from_scratch = TRUE,
    known_times = known_times
  )
  balance
}
}
})

## End(Not run)

```

predict_workers	<i>Predict the load balancing of the next call to <code>make()</code> for non-staged parallel backends. Maturing</i>
-----------------	---

Description

Take the past recorded runtimes times from `build_times()` and use them to predict how the targets will be distributed among the available workers in the next `make()`. Predictions only include the time it takes to run the targets, not overhead/preprocessing from drake itself.

Usage

```

predict_workers(config, targets = NULL, from_scratch = FALSE,
  targets_only = NULL, jobs = 1, known_times = numeric(0),
  default_time = 0, warn = TRUE)

```

Arguments

config	Optional internal runtime parameter list of produced by both <code>make()</code> and <code>drake_config()</code> .
targets	Character vector, names of targets. Predict the runtime of building these targets plus dependencies. Defaults to all targets.
from_scratch	Logical, whether to predict a <code>make()</code> build from scratch or to take into account the fact that some targets may be already up to date and therefore skipped.
targets_only	Deprecated.
jobs	The <code>jobs</code> argument of your next planned <code>make()</code> . How many targets to do you plan to have running simultaneously?
known_times	A named numeric vector with targets/imports as names and values as hypothetical runtimes in seconds. Use this argument to overwrite any of the existing build times or the <code>default_time</code> .
default_time	Number of seconds to assume for any target or import with no recorded runtime (from <code>build_times()</code>) or anything in <code>known_times</code> .
warn	Logical, whether to warn the user about any targets with no available runtime, either in <code>known_times</code> or <code>build_times()</code> . The times for these targets default to <code>default_time</code> .

Value

A data frame showing one likely arrangement of targets assigned to parallel workers.

See Also

`predict_runtime()`, `build_times()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    config <- drake_config(my_plan)
    known_times <- rep(7200, nrow(my_plan))
    names(known_times) <- my_plan$target
    known_times
    # Predict the runtime
    if (requireNamespace("lubridate", quietly = TRUE)) {
      predict_runtime(
        config = config,
        jobs = 7,
        from_scratch = TRUE,
        known_times = known_times
      )
      predict_runtime(
        config,
        jobs = 8,
        from_scratch = TRUE,
        known_times = known_times
      )
      balance <- predict_workers(
        config,
        jobs = 7,
        from_scratch = TRUE,
        known_times = known_times
      )
      balance
    }
  }
})

## End(Not run)
```

progress

Get the build progress of your targets **Maturing**

Description

Objects that drake imported, built, or attempted to build are listed as "done" or "running". Skipped objects are not listed.

Usage

```
progress(..., list = character(0), no_imported_objects = NULL,
  path = NULL, search = NULL, cache = drake::drake_cache(path = path,
  verbose = verbose), verbose = 1L, jobs = 1, progress = NULL)
```

Arguments

...	Objects to load from the cache, as names (unquoted) or character strings (quoted). If the <code>tidyselect</code> package is installed, you can also supply dplyr-style <code>tidyselect</code> commands such as <code>starts_with()</code> , <code>ends_with()</code> , and <code>one_of()</code> .
<code>list</code>	Character vector naming objects to be loaded from the cache. Similar to the <code>list</code> argument of <code>remove()</code> .
<code>no_imported_objects</code>	Logical, whether to only return information about imported files and targets with commands (i.e. whether to ignore imported objects that are not files).
<code>path</code>	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
<code>search</code>	Deprecated.
<code>cache</code>	drake cache. See <code>new_cache()</code> . If supplied, <code>path</code> is ignored.
<code>verbose</code>	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
<code>jobs</code>	Number of jobs/workers for parallel processing.
<code>progress</code>	Character vector for filtering the build progress results. Defaults to <code>NULL</code> (no filtering) to report progress of all objects. Supported filters are "done", "running", and "failed".

Value

The build progress of each target reached by the current `make()` so far.

See Also

`diagnose()`, `drake_get_session_info()`, `cached()`, `readd()`, `drake_plan()`, `make()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    # Watch the changing progress() as make() is running.
    progress() # List all the targets reached so far.
    progress(small, large) # Just see the progress of some targets.
    progress(list = c("small", "large")) # Same as above.
  }
})

## End(Not run)
```

readd

*Read and return a drake target/import from the cache. **Maturing*****Description**

`readd()` returns an object from the cache, and `loadadd()` loads one or more objects from the cache into your environment or session. These objects are usually targets built by `make()`.

Usage

```
readd(target, character_only = FALSE, path = NULL, search = NULL,
      cache = drake::drake_cache(path = path, verbose = verbose),
      namespace = NULL, verbose = 1L, show_source = FALSE)

loadadd(..., list = character(0), imported_only = NULL, path = NULL,
        search = NULL, cache = drake::drake_cache(path = path, verbose =
        verbose), namespace = NULL, envir = parent.frame(), jobs = 1,
        verbose = 1L, deps = FALSE, lazy = "eager", graph = NULL,
        replace = TRUE, show_source = FALSE, tidyselect = !deps,
        config = NULL)
```

Arguments

target	If <code>character_only</code> is TRUE, then <code>target</code> is a character string naming the object to read. Otherwise, <code>target</code> is an unquoted symbol with the name of the object.
character_only	Logical, whether name should be treated as a character or a symbol (just like <code>character.only</code> in <code>library()</code>).
path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or NULL.
search	Deprecated.
cache	drake cache. See <code>new_cache()</code> . If supplied, <code>path</code> is ignored.
namespace	Optional character string, name of the storrr namespace to read from.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
show_source	Logical, option to show the command that produced the target or indicate that the object was imported (using <code>show_source()</code>).
...	Targets to load from the cache: as names (symbols) or character strings. If the <code>tidyselect</code> package is installed, you can also supply dplyr-style <code>tidyselect</code> commands such as <code>starts_with()</code> , <code>ends_with()</code> , and <code>one_of()</code> .
list	Character vector naming targets to be loaded from the cache. Similar to the <code>list</code> argument of <code>remove()</code> .
imported_only	Logical, deprecated.
envir	Environment to load objects into. Defaults to the calling environment (current workspace).

jobs	Number of parallel jobs for loading objects. On non-Windows systems, the loading process for multiple objects can be lightly parallelized via <code>parallel::mclapply()</code> . Just set jobs to be an integer greater than 1. On Windows, jobs is automatically demoted to 1.
deps	Logical, whether to load any cached dependencies of the targets instead of the targets themselves. Important note: <code>deps = TRUE</code> disables <code>tidyselect</code> functionality. For example, <code>loadadd(starts_with("model_"), config = config, deps = TRUE)</code> does not work. For the selection mechanism to work, the <code>model_*</code> targets need to already be in the cache, which is not always the case when you are debugging your projects. To help drake understand what you mean, you must name the targets <i>explicitly</i> when <code>deps</code> is <code>TRUE</code> , e.g. <code>loadadd(model_A, model_B, config = config, deps = TRUE)</code> .
lazy	Either a string or a logical. Choices: <ul style="list-style-type: none"> • "eager": no lazy loading. The target is loaded right away with <code>assign()</code>. • "promise": lazy loading with <code>delayedAssign()</code> • "bind": lazy loading with active bindings: <code>bindr::populate_env()</code>. • <code>TRUE</code>: same as "promise". • <code>FALSE</code>: same as "eager".
graph	Deprecated.
replace	Logical. If <code>FALSE</code> , items already in your environment will not be replaced.
tidyselect	Logical, whether to enable <code>tidyselect</code> expressions in ... like <code>starts_with("prefix")</code> and <code>ends_with("suffix")</code> .
config	Optional <code>drake_config()</code> object. You should supply one if <code>deps</code> is <code>TRUE</code> .

Details

There are two uses for the `loadadd()` and `readd()` functions:

1. Exploring the results outside the drake/make() pipeline. When you call `make()` to run your project, drake puts the targets in a cache, usually a folder called `.drake`. You may want to inspect the targets afterwards, possibly in an interactive R session. However, the files in the `.drake` folder are organized in a special format created by the `storr` package, which is not exactly human-readable. To retrieve a target for manual viewing, use `readd()`. To load one or more targets into your session, use `loadadd()`.
2. In knitr / R Markdown reports. You can borrow drake targets in your active code chunks if you have the right calls to `loadadd()` and `readd()`. These reports can either run outside the drake pipeline, or better yet, as part of the pipeline itself. If you call `knitr_in("your_report.Rmd")` inside a `drake_plan()` command, then `make()` will scan "your_report.Rmd" for calls to `loadadd()` and `readd()` in active code chunks, and then treat those loaded targets as dependencies. That way, `make()` will automatically (re)run the report if those dependencies change.
3. If you are using `make(memory_strategy = "none")` or `make(memory_strategy = "unload")`, `loadadd()` and `readd()` can manually load dependencies into memory for the target that is being built. If you do this, you must carefully inspect `deps_target()` and `vis_drake_graph()` before running `make()` to be sure the dependency relationships among targets are correct. If you do not wish to incur extra dependencies with `loadadd()` or `readd()`, you will need to use `ignore()`, e.g. `drake_plan(x = 1, y = ignore(readd(x)))` or `drake_plan(x = 1, y = readd(ignore("x"), character_only = TRUE))`. Compare those plans to `drake_plan(x = 1, y = readd(x))` and `drake_plan(x = 1, y = readd("x", character_only = TRUE))` using `vis_drake_graph()` and `deps_target()`.

Value

The cached value of the target.

See Also

[cached\(\)](#), [drake_plan\(\)](#), [make\(\)](#)
[cached\(\)](#), [drake_plan\(\)](#), [make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    readd(reg1) # Return imported object 'reg1' from the cache.
    readd(small) # Return targets 'small' from the cache.
    readd("large", character_only = TRUE) # Return 'large' from the cache.
    # For external files, only the fingerprint/hash is stored.
    readd(file_store("report.md"), character_only = TRUE)
  }
})

## End(Not run)

## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the projects, build the targets.
    config <- drake_config(my_plan)
    loadd(small) # Load target 'small' into your workspace.
    small
    # For many targets, you can parallelize loadd()
    # using the 'jobs' argument.
    loadd(list = c("small", "large"), jobs = 2)
    ls()
    # Load the dependencies of the target, coef_regression2_small
    loadd(coef_regression2_small, deps = TRUE, config = config)
    ls()
    # Load all the targets listed in the workflow plan
    # of the previous `make()`.
    # If you do not supply any target names, `loadd()` loads all the targets.
    # Be sure your computer has enough memory.
    loadd()
    ls()
  }
})

## End(Not run)
```


Description

When a project is created with `make()` or `drake_config()`, the project's pseudo-random number generator seed is cached. Then, unless the cache is destroyed, the seeds of all the targets will deterministically depend on this one central seed. That way, reproducibility is protected, even under randomness.

Usage

```
read_drake_seed(path = NULL, search = NULL, cache = NULL,
  verbose = 1L)
```

Arguments

<code>path</code>	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
<code>search</code>	Deprecated.
<code>cache</code>	drake cache. See <code>new_cache()</code> . If supplied, <code>path</code> is ignored.
<code>verbose</code>	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.

Value

An integer vector.

Examples

```
## Not run:
isolate_example("contain side effects", {
  cache <- storr::storr_environment() # Just for the examples.
  my_plan <- drake_plan(
    target1 = sqrt(1234),
    target2 = sample.int(n = 12, size = 1) + target1
  )
  tmp <- sample.int(1) # Needed to get a .Random.seed, but not for drake.
  digest::digest(.Random.seed) # Fingerprint of the current R session's seed.
  make(my_plan, cache = cache) # Run the project, build the targets.
  digest::digest(.Random.seed) # Your session's seed did not change.
  # drake uses a hard-coded seed if you do not supply one.
  read_drake_seed(cache = cache)
  readd(target2, cache = cache) # Randomly-generated target data.
  clean(target2, cache = cache) # Oops, I removed the data!
  tmp <- sample.int(1) # Maybe the R session's seed also changed.
  make(my_plan, cache = cache) # Rebuild target2.
  # Same as before:
  read_drake_seed(cache = cache)
  readd(target2, cache = cache)
  # You can also supply a seed.
  # If your project already exists, it must agree with the project's
  # preexisting seed (default: 0)
  clean(target2, cache = cache)
  make(my_plan, cache = cache, seed = 0)
  read_drake_seed(cache = cache)
  readd(target2, cache = cache)
```

```
# If you want to supply a different seed than 0,
# you need to destroy the cache and start over first.
clean(destroy = TRUE, cache = cache)
cache <- storr::storr_environment() # Just for the examples.
make(my_plan, cache = cache, seed = 1234)
read_drake_seed(cache = cache)
readd(target2, cache = cache)
})

## End(Not run)
```

recoverable

*List the most upstream recoverable outdated targets. **Experimental***

Description

Only shows the most upstream updated targets. Whether downstream targets are recoverable depends on the eventual values of the upstream targets in the next [make\(\)](#).

Usage

```
recoverable(config, make_imports = TRUE, do_prework = TRUE)
```

Arguments

config	Optional internal runtime parameter list produced with drake_config() . You must use a fresh config argument with an up-to-date dependency graph that was never modified by hand. If needed, rerun drake_config() early and often. See the details in the help file for drake_config() .
make_imports	Logical, whether to make the imports first. Set to FALSE to save some time and risk obsolete output.
do_prework	Whether to do the prework normally supplied to make() .

Value

Character vector of the names of recoverable targets.

Recovery

`make(recover = TRUE, recoverable = TRUE)` powers automated data recovery. The default of `recover` is FALSE because

1. Automated data recovery is still experimental.
2. It has reproducibility issues. Targets recovered from the distant past may have been generated with earlier versions of R and earlier package environments that no longer exist.

How it works: if `recover` is TRUE, drake tries to salvage old target values from the cache instead of running commands from the plan. A target is recoverable if

1. There is an old value somewhere in the cache that shares the command, dependencies, etc. of the target about to be built.
2. The old value was generated with `make(recoverable = TRUE)`.

If both conditions are met, drake will

1. Assign the most recently-generated admissible data to the target, and
2. skip the target's command.

See Also

[r_recoverable\(\)](#), [r_outdated\(\)](#), [drake_config\(\)](#), [missed\(\)](#), [drake_plan\(\)](#), [make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    config <- drake_config(my_plan)
    make(my_plan)
    clean()
    outdated(config) # Which targets are outdated?
    recoverable(config) # Which of these are recoverable and upstream?
    # The report still builds because clean() removes report.md,
    # but make() recovers the rest.
    make(my_plan, recover = TRUE)
    # When was the *recovered* small data actually built (first stored)?
    # (Was I using a different version of R back then?)
    diagnose(small)$date
    # If you set the same seed as before, you can even
    # rename targets without having to build them again.
    # For an example, see
    # the "Reproducible data recovery and renaming" section of
    # https://github.com/ropensci/drake/blob/master/README.md.
  }
})

## End(Not run)
```

render_drake_ggraph	<i>Visualize the workflow with</i> drake_graph_info() <i>output.</i> Stable	ggplot2/ggraph using
---------------------	---	----------------------

Description

This function requires packages ggplot2 and ggraph. Install them with `install.packages(c("ggplot2", "ggraph"))`

Usage

```
render_drake_ggraph(graph_info, main = graph_info$default_title,
  label_nodes = FALSE, transparency = TRUE)
```

Arguments

graph_info	List of data frames generated by <code>drake_graph_info()</code> . There should be 3 data frames: nodes, edges, and legend_nodes.
main	Character string, title of the graph.
label_nodes	Logical, whether to label the nodes. If FALSE, the graph will not have any text next to the nodes, which is recommended for large graphs with lots of targets.
transparency	Logical, whether to allow transparency in the rendered graph. Set to FALSE if you get warnings like "semi-transparency is not supported on this device".

Value

A ggplot2 object, which you can modify with more layers, show with `plot()`, or save as a file with `ggsave()`.

See Also

[vis_drake_graph\(\)](#), [sankey_drake_graph\(\)](#), [drake_ggraph\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  load_mtcars_example() # Get the code with drake_example("mtcars").
  if (requireNamespace("ggraph", quietly = TRUE)) {
    # Instead of jumping right to vis_drake_graph(), get the data frames
    # of nodes, edges, and legend nodes.
    config <- drake_config(my_plan) # Internal configuration list
    drake_ggraph(config) # Jump straight to the static graph.
    # Get the node and edge info that vis_drake_graph() just plotted:
    graph <- drake_graph_info(config)
    render_drake_ggraph(graph)
  }
})

## End(Not run)
```

render_drake_graph	<i>Render a visualization using the data frames generated by <code>drake_graph_info()</code>. Maturing</i>
--------------------	---

Description

This function is called inside `vis_drake_graph()`, which typical users call more often.

Usage

```
render_drake_graph(graph_info, file = character(0), layout = NULL,
  direction = NULL, hover = TRUE, main = graph_info$default_title,
  selfcontained = FALSE, navigationButtons = TRUE, ncol_legend = 1,
  collapse = TRUE, ...)
```

Arguments

graph_info	List of data frames generated by <code>drake_graph_info()</code> . There should be 3 data frames: nodes, edges, and legend_nodes.
file	Name of a file to save the graph. If NULL or character(0), no file is saved and the graph is rendered and displayed within R. If the file ends in a .png, .jpg, .jpeg, or .pdf extension, then a static image will be saved. In this case, the webshot package and PhantomJS are required: <code>install.packages("webshot"); webshot::install_phantomjs()</code> . If the file does not end in a .png, .jpg, .jpeg, or .pdf extension, an HTML file will be saved, and you can open the interactive graph using a web browser.
layout	Deprecated.
direction	Deprecated.
hover	Logical, whether to show the command that generated the target when you hover over a node with the mouse. For imports, the label does not change with hovering.
main	Character string, title of the graph.
selfcontained	Logical, whether to save the file as a self-contained HTML file (with external resources base64 encoded) or a file with external resources placed in an adjacent directory. If TRUE, pandoc is required. The selfcontained argument only applies to HTML files. In other words, if file is a PNG, PDF, or JPEG file, for instance, the point is moot.
navigationButtons	Logical, whether to add navigation buttons with <code>visNetwork::visInteraction(navigationButtons = TRUE)</code>
ncol_legend	Number of columns in the legend nodes. To remove the legend entirely, set ncol_legend to NULL or 0.
collapse	Logical, whether to allow nodes to collapse if you double click on them. Analogous to <code>visNetwork::visOptions(collapse = TRUE)</code> or <code>visNetwork::visOptions(collapse = TRUE)</code> .
...	Arguments passed to <code>visNetwork()</code> .

Value

A `visNetwork` graph.

See Also

`vis_drake_graph()`, `sankey_drake_graph()`, `drake_ggraph()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    if (requireNamespace("visNetwork", quietly = TRUE)) {
      # Instead of jumping right to vis_drake_graph(), get the data frames
      # of nodes, edges, and legend nodes.
      config <- drake_config(my_plan) # Internal configuration list
      vis_drake_graph(config) # Jump straight to the interactive graph.
```

```
# Get the node and edge info that vis_drake_graph() just plotted:
graph <- drake_graph_info(config)
# You can pass the data frames right to render_drake_graph()
# (as in vis_drake_graph()) or you can create
# your own custom visNetwork graph.
render_drake_graph(graph, width = '100%') # Width is passed to visNetwork.
}
}
})

## End(Not run)
```

render_sankey_drake_graph

Render a Sankey diagram from `drake_graph_info()`. Stable

Description

This function is called inside `sankey_drake_graph()`, which typical users call more often. A legend is unfortunately unavailable for the graph itself (<https://github.com/christophergandrud/networkD3/issues/240>) but you can see what all the colors mean with `visNetwork::visNetwork(drake::legend_no`

Usage

```
render_sankey_drake_graph(graph_info, file = character(0),
  selfcontained = FALSE, ...)
```

Arguments

<code>graph_info</code>	List of data frames generated by <code>drake_graph_info()</code> . There should be 3 data frames: nodes, edges, and legend_nodes.
<code>file</code>	Name of a file to save the graph. If <code>NULL</code> or <code>character(0)</code> , no file is saved and the graph is rendered and displayed within R. If the file ends in a <code>.png</code> , <code>.jpg</code> , <code>.jpeg</code> , or <code>.pdf</code> extension, then a static image will be saved. In this case, the <code>webshot</code> package and <code>PhantomJS</code> are required: <code>install.packages("webshot"); webshot::install_phantomjs()</code> . If the file does not end in a <code>.png</code> , <code>.jpg</code> , <code>.jpeg</code> , or <code>.pdf</code> extension, an HTML file will be saved, and you can open the interactive graph using a web browser.
<code>selfcontained</code>	Logical, whether to save the file as a self-contained HTML file (with external resources base64 encoded) or a file with external resources placed in an adjacent directory. If <code>TRUE</code> , <code>pandoc</code> is required.
<code>...</code>	Arguments passed to <code>networkD3::sankeyNetwork()</code> .

Value

A `visNetwork` graph.

See Also

[sankey_drake_graph\(\)](#), [vis_drake_graph\(\)](#), [drake_ggraph\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  load_mtcars_example() # Get the code with drake_example("mtcars").
  if (suppressWarnings(require("knitr"))){
    if (requireNamespace("networkD3", quietly = TRUE)) {
      if (requireNamespace("visNetwork", quietly = TRUE)) {
        # Instead of jumping right to sankey_drake_graph(), get the data frames
        # of nodes, edges, and legend nodes.
        config <- drake_config(my_plan) # Internal configuration list
        sankey_drake_graph(config) # Jump straight to the interactive graph.
        # Show the legend separately.
        visNetwork::visNetwork(nodes = drake::legend_nodes())
        # Get the node and edge info that sankey_drake_graph() just plotted:
        graph <- drake_graph_info(config)
        # You can pass the data frames right to render_sankey_drake_graph()
        # (as in sankey_drake_graph()) or you can create
        # your own custom visNetwork graph.
        render_sankey_drake_graph(graph, width = '100%')
      }
    }
  }
})

## End(Not run)
```

render_text_drake_graph

*Show a workflow graph as text in your terminal window using
drake_graph_info() output. **Stable***

Description

This function is called inside [text_drake_graph\(\)](#), which typical users call more often. See [?text_drake_graph](#) for details.

Usage

```
render_text_drake_graph(graph_info, nchar = 1L, print = TRUE)
```

Arguments

graph_info	List of data frames generated by drake_graph_info() . There should be 3 data frames: nodes, edges, and legend_nodes.
nchar	For each node, maximum number of characters of the node label to show. Can be 0, in which case each node is a colored box instead of a node label. Caution: nchar > 0 will mess with the layout.
print	Logical. If TRUE, the graph will print to the console via message() . If FALSE, nothing is printed. However, you still have the visualization because text_drake_graph() and render_text_drake_graph() still invisibly return a character string that you can print yourself with message() .

Value

The lines of text in the visualization.

See Also

[text_drake_graph\(\)](#), [vis_drake_graph\(\)](#), [sankey_drake_graph\(\)](#), [drake_ggraph\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    pkgs <- requireNamespace("txtplot", quietly = TRUE) &&
      requireNamespace("visNetwork", quietly = TRUE)
    if (pkgs) {
      # Instead of jumping right to vis_drake_graph(), get the data frames
      # of nodes, edges, and legend nodes.
      config <- drake_config(my_plan) # Internal configuration list
      text_drake_graph(config) # Jump straight to the interactive graph.
      # Get the node and edge info that vis_drake_graph() just plotted:
      graph <- drake_graph_info(config)
      # You can pass the data frames right to render_text_drake_graph().
      render_text_drake_graph(graph)
    }
  }
})

## End(Not run)
```

rescue_cache

*Try to repair a drake cache that is prone to throwing storrr-related errors. **Questioning***

Description

Sometimes, storrr caches may have dangling orphaned files that prevent you from loading or cleaning. This function tries to remove those files so you can use the cache normally again.

Usage

```
rescue_cache(targets = NULL, path = NULL, search = NULL,
  verbose = 1L, force = FALSE, cache = drake::drake_cache(path =
  path, verbose = verbose), jobs = 1, garbage_collection = FALSE)
```

Arguments

targets	Character vector, names of the targets to rescue. As with many other drake utility functions, the word target is defined generally in this case, encompassing imports as well as true targets. If targets is NULL, everything in the cache is rescued.
---------	--

path	Character. Set path to the path of a <code>storr::storr_rds()</code> cache to retrieve a specific cache generated by <code>storr::storr_rds()</code> or <code>drake::new_cache()</code> . If the path argument is <code>NULL</code> , <code>drake_cache()</code> searches up through parent directories to find a folder called <code>.drake/</code> .
search	Deprecated.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.
force	Deprecated.
cache	A <code>storr</code> cache object.
jobs	Number of jobs for light parallelism (disabled on Windows).
garbage_collection	Logical, whether to do garbage collection as a final step. See <code>drake_gc()</code> and <code>clean()</code> for details.

Value

Nothing.

See Also

`drake_cache()`, `cached()`, `drake_gc()`, `clean()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build targets. This creates the cache.
    # Remove dangling cache files that could cause errors.
    rescue_cache(jobs = 2)
    # Alternatively, just rescue targets 'small' and 'large'.
    # Rescuing specific targets is usually faster.
    rescue_cache(targets = c("small", "large"))
  }
})

## End(Not run)
```

running

List running targets. **Maturing**

Description

List the targets that either (1) are currently being built during a call to `make()`, or (2) if `make()` was interrupted, the targets that were running at the time.

Usage

```
running(path = NULL, search = NULL, cache = drake::drake_cache(path =
  path, verbose = verbose), verbose = 1L)
```

Arguments

path	Path to a drake cache (usually a hidden <code>.drake/</code> folder) or <code>NULL</code> .
search	Deprecated.
cache	drake cache. See new_cache() . If supplied, path is ignored.
verbose	Integer, control printing to the console/terminal. <ul style="list-style-type: none"> • 0: print nothing. • 1: print targets, retries, and failures. • 2: also show a spinner when preprocessing tasks are underway.

Value

A character vector of target names.

See Also

[failed\(\)](#), [make\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    make(my_plan) # Run the project, build the targets.
    running() # Everything should be done.
    # nolint start
    # Run make() in one R session...
    # slow_plan <- drake_plan(x = Sys.sleep(2))
    # make(slow_plan)
    # and see the progress in another session.
    # running()
    # nolint end
  }
})

## End(Not run)
```

r_make

Launch a drake function in a fresh new R process **Maturing**

Description

The `r_*`() functions, such as `r_make()`, enhance reproducibility by launching a drake function in a separate R process.

Usage

```

r_make(source = NULL, r_fn = NULL, r_args = list())

r_drake_build(target, character_only = FALSE, ..., source = NULL,
  r_fn = NULL, r_args = list())

r_outdated(..., source = NULL, r_fn = NULL, r_args = list())

r_recoverable(..., source = NULL, r_fn = NULL, r_args = list())

r_missed(..., source = NULL, r_fn = NULL, r_args = list())

r_deps_target(target, character_only = FALSE, ..., source = NULL,
  r_fn = NULL, r_args = list())

r_drake_graph_info(..., source = NULL, r_fn = NULL, r_args = list())

r_vis_drake_graph(..., source = NULL, r_fn = NULL, r_args = list())

r_sankey_drake_graph(..., source = NULL, r_fn = NULL,
  r_args = list())

r_drake_ggraph(..., source = NULL, r_fn = NULL, r_args = list())

r_text_drake_graph(..., source = NULL, r_fn = NULL, r_args = list())

r_predict_runtime(..., source = NULL, r_fn = NULL, r_args = list())

r_predict_workers(..., source = NULL, r_fn = NULL, r_args = list())

```

Arguments

source	Path to an R script file that loads packages, functions, etc. and returns a <code>drake_config()</code> object. There are 3 ways to set this path. <ol style="list-style-type: none"> 1. Pass an explicit file path. 2. Call <code>options(drake_source = "path_to_your_script.R")</code>. 3. Just create a file called <code>"_drake.R"</code> in your working directory and supply nothing to source.
r_fn	A callr function such as <code>callr::r</code> or <code>callr::r_bg</code> . Example: <code>r_make(r_fn = callr::r)</code> .
r_args	List of arguments to <code>r_fn</code> , not including <code>func</code> or <code>args</code> . Example: <code>r_make(r_fn = callr::r_bg, r_args = list(stdout = "stdout.log"))</code> .
target	Name of the target.
character_only	Logical, whether name should be treated as a character or a symbol (just like <code>character.only</code> in <code>library()</code>).
...	Arguments to the inner function. For example, if you want to call <code>r_vis_drake_graph()</code> , the inner function is <code>vis_drake_graph()</code> , and <code>selfcontained</code> is an example argument you could supply to the ellipsis.

Details

drake searches your environment to detect dependencies, so functions like `make()`, `outdated()`, etc. are designed to run in fresh clean R sessions. Wrappers `r_make()`, `r_outdated()`, etc. run reproducibly even if your current R session is old and stale.

`r_outdated()` runs the four steps below. `r_make()` etc. are similar.

1. Launch a new `callr::r()` session.
2. In that fresh session, run the R script from the `source` argument. This script loads packages, functions, global options, etc. and returns a `drake_config()` object.
3. In that same session, run `outdated()` with the `config` argument from step 2.
4. Return the result back to master process (e.g. your interactive R session).

Recovery

`make(recover = TRUE, recoverable = TRUE)` powers automated data recovery. The default of `recover` is `FALSE` because

1. Automated data recovery is still experimental.
2. It has reproducibility issues. Targets recovered from the distant past may have been generated with earlier versions of R and earlier package environments that no longer exist.

How it works: if `recover` is `TRUE`, drake tries to salvage old target values from the cache instead of running commands from the plan. A target is recoverable if

1. There is an old value somewhere in the cache that shares the command, dependencies, etc. of the target about to be built.
2. The old value was generated with `make(recoverable = TRUE)`.

If both conditions are met, drake will

1. Assign the most recently-generated admissible data to the target, and
2. skip the target's command.

See Also

`make()`

Examples

```
## Not run:
isolate_example("quarantine side effects", {
  writeLines(
    c(
      "library(drake)",
      "load_mtcars_example()",
      "drake_config(my_plan)"
    ),
    "_drake.R" # default value of the `source` argument
  )
  cat(readLines("_drake.R"), sep = "\n")
  r_outdated()
  r_make()
  r_outdated()
})
```

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

sankey_drake_graph	Show a Sankey graph of your drake project. Stable
--------------------	--

Description

To save time for repeated plotting, this function is divided into `drake_graph_info()` and `render_sankey_drake_graph`. A legend is unfortunately unavailable for the graph itself (<https://github.com/christophergandrud/networkD3/issues/240>) but you can see what all the colors mean with `visNetwork::visNetwork(drake::legend_no`

Usage

```
sankey_drake_graph(config, file = character(0), selfcontained = FALSE,
  build_times = "build", digits = 3, targets_only = FALSE,
  from = NULL, mode = c("out", "in", "all"), order = NULL,
  subset = NULL, make_imports = TRUE, from_scratch = FALSE,
  group = NULL, clusters = NULL, show_output_files = TRUE, ...)
```

Arguments

config	A <code>drake_config()</code> configuration list. You can get one as a return value from <code>make()</code> as well.
file	Name of a file to save the graph. If <code>NULL</code> or <code>character(0)</code> , no file is saved and the graph is rendered and displayed within R. If the file ends in a <code>.png</code> , <code>.jpg</code> , <code>.jpeg</code> , or <code>.pdf</code> extension, then a static image will be saved. In this case, the <code>webshot</code> package and <code>PhantomJS</code> are required: <code>install.packages("webshot"); webshot::install_phantomjs()</code> . If the file does not end in a <code>.png</code> , <code>.jpg</code> , <code>.jpeg</code> , or <code>.pdf</code> extension, an HTML file will be saved, and you can open the interactive graph using a web browser.
selfcontained	Logical, whether to save the file as a self-contained HTML file (with external resources base64 encoded) or a file with external resources placed in an adjacent directory. If <code>TRUE</code> , <code>pandoc</code> is required.
build_times	Character string or logical. If character, the choices are 1. "build": runtime of the command plus the time it take to store the target or import. 2. "command": just the runtime of the command. 3. "none": no build times. If logical, <code>build_times</code> selects whether to show the times from 'build_times(..., type = "build")' or use no build times at all. See <code>build_times()</code> for details.
digits	Number of digits for rounding the build times
targets_only	Logical, whether to skip the imports and only include the targets in the workflow plan.
from	Optional collection of target/import names. If <code>from</code> is nonempty, the graph will restrict itself to a neighborhood of <code>from</code> . Control the neighborhood with <code>mode</code> and <code>order</code> .
mode	Which direction to branch out in the graph to create a neighborhood around <code>from</code> . Use "in" to go upstream, "out" to go downstream, and "all" to go both ways and disregard edge direction altogether.

order	How far to branch out to create a neighborhood around from. Defaults to as far as possible. If a target is in the neighborhood, then so are all of its custom <code>file_out()</code> files if <code>show_output_files</code> is TRUE. That means the actual graph order may be slightly greater than you might expect, but this ensures consistency between <code>show_output_files = TRUE</code> and <code>show_output_files = FALSE</code> .
subset	Optional character vector. Subset of targets/imports to display in the graph. Applied after from, mode, and order. Be advised: edges are only kept for adjacent nodes in subset. If you do not select all the intermediate nodes, edges will drop from the graph.
make_imports	Logical, whether to make the imports first. Set to FALSE to increase speed and risk using obsolete information.
from_scratch	Logical, whether to assume all the targets will be made from scratch on the next <code>make()</code> . Makes all targets outdated, but keeps information about build progress in previous <code>make()</code> s.
group	Optional character scalar, name of the column used to group nodes into columns. All the columns names of your original drake plan are choices. The other choices (such as "status") are column names in the nodes . To group nodes into clusters in the graph, you must also supply the clusters argument.
clusters	Optional character vector of values to cluster on. These values must be elements of the column of the nodes data frame that you specify in the group argument to <code>drake_graph_info()</code> .
show_output_files	Logical, whether to include <code>file_out()</code> files in the graph.
...	Arguments passed to <code>networkD3::sankeyNetwork()</code> .

Value

A `visNetwork` graph.

See Also

`render_sankey_drake_graph()`, `vis_drake_graph()`, `drake_ggraph()`, `text_drake_graph()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    config <- drake_config(my_plan)
    if (requireNamespace("networkD3", quietly = TRUE)) {
      if (requireNamespace("visNetwork", quietly = TRUE)) {
        # Plot the network graph representation of the workflow.
        sankey_drake_graph(config, width = '100%')
        # Show the legend separately.
        visNetwork::visNetwork(nodes = drake::legend_nodes())
        make(my_plan) # Run the project, build the targets.
        sankey_drake_graph(config) # The black nodes from before are now green.
        # Plot a subgraph of the workflow.
        sankey_drake_graph(config, from = c("small", "reg2"))
      }
    }
  }
}
```

```
  })

  ## End(Not run)
```

show_source	<i>Show how a target/import was produced. Stable</i>
-------------	---

Description

Show the command that produced a target or indicate that the object or file was imported.

Usage

```
show_source(target, config, character_only = FALSE)
```

Arguments

target	Symbol denoting the target or import or a character vector if character_only is TRUE.
config	A drake_config() list.
character_only	Logical, whether to interpret target as a symbol (FALSE) or character vector (TRUE).

Examples

```
## Not run:
isolate_example("contain side effects", {
  plan <- drake_plan(x = sample.int(15))
  cache <- storr::storr_environment() # custom in-memory cache
  make(plan, cache = cache)
  config <- drake_config(plan, cache = cache)
  show_source(x, config)
})

## End(Not run)
```

split	<i>Define a target for each subset of data Maturing</i>
-------	--

Description

Similar group_map(), from dplyr, except it defines new targets in drake.

Usage

```
split(..., .id, .tag_in, .tag_out)
```

Arguments

<code>...</code>	Grouping variables. New grouping variables must be supplied with their names and values, existing grouping variables can be given as symbols without any values assigned.
<code>.id</code>	Symbol or vector of symbols naming grouping variables to incorporate into target names. Useful for creating short target names. Set <code>.id = FALSE</code> to use integer indices as target name suffixes.
<code>.tag_in</code>	A symbol or vector of symbols. Tags assign targets to grouping variables. Use <code>.tag_in</code> to assign <i>untransformed</i> targets to grouping variables.
<code>.tag_out</code>	Just like <code>.tag_in</code> , except that <code>.tag_out</code> assigns <i>transformed</i> targets to grouping variables.

Details

Only valid within a call to `target()` in `drake_plan()`. See the examples below.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), # nolint supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), # nolint or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`map`, `cross`, `combine`, `drake_plan`, `target`, `drake_slice`

Examples

```
plan <- drake_plan(
  analysis = target(
    analyze(data),
    transform = split(data, slices = 3L, margin = 1L, drop = FALSE)
  )
)
print(plan)
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}
```

target

Customize a target in `drake_plan()`. **Maturing**

Description

Must be called inside `drake_plan()`. Invalid otherwise.

Usage

```
target(command = NULL, ...)
```

Arguments

command	The command to build the target.
...	Optional columns of the plan for a given target. See the Columns section of this help file for a selection of special columns that drake understands.

Value

A one-row workflow plan data frame with the named arguments as columns.

Columns

`drake_plan()` creates a special data frame. At minimum, that data frame must have columns `target` and `command` with the target names and the R code chunks to build them, respectively.

You can add custom columns yourself, either with `target()` (e.g. `drake_plan(y = target(f(x), transform = map(c(1, 2)), format = "fst"))`) or by appending columns post-hoc (e.g. `plan$col <- vals`).

Some of these custom columns are special. They are optional, but drake looks for them at various points in the workflow.

- `transform`: a call to `map()`, `split()`, `cross()`, or `combine()` to create and manipulate large collections of targets. Details: (<https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>). # nolint
- `format`: set a storage format to save big targets more efficiently. Most formats are faster than ordinary storage, and they consume far less memory. Available formats:
 - `"fst"`: save big data frames fast. The target must be a data frame, and you must have the `fst` package installed.
 - `"keras"`: save Keras models as HDF5 files. Requires the `keras` package.

- "rds": save any object. This is similar to the default storage except we avoid creating a serialized copy of the entire target in memory. Requires R \geq 3.5.0 so drake can use ALTREP.
- trigger: rule to decide whether a target needs to run. It is recommended that you define this one with `target()`. Details: <https://ropenscilabs.github.io/drake-manual/triggers.html>.
- hpc: logical values (TRUE/FALSE/NA) whether to send each target to parallel workers. Visit <https://ropenscilabs.github.io/drake-manual/hpc.html#selectivity> to learn more.
- resources: target-specific lists of resources for a computing cluster. See <https://ropenscilabs.github.io/drake-manual/hpc.html#advanced-options> for details.
- elapsed and cpu: number of seconds to wait for the target to build before timing out (elapsed for elapsed time and cpu for CPU time).
- retries: number of times to retry building a target in the event of an error.
- seed: an optional pseudo-random number generator (RNG) seed for each target. drake usually comes up with its own unique reproducible target-specific seeds using the global seed (the seed argument to `make()` and `drake_config()`) and the target names, but you can overwrite these automatic seeds. NA entries default back to drake's automatic seeds.

Keywords

`drake_plan()` understands special keyword functions for your commands. With the exception of `target()`, each one is a proper function with its own help file.

- `target()`: give the target more than just a command. Using `target()`, you can apply a transformation (examples: <https://ropenscilabs.github.io/drake-manual/plans.html#large-plans>), # nolint supply a trigger (<https://ropenscilabs.github.io/drake-manual/triggers.html>), # nolint or set any number of custom columns.
- `map()`: create multiple targets. Called inside `target()`.
- `split()`: create a target for each subset of data. Called inside `target()`.
- `cross()`: create a target for each combination of values. Called inside `target()`.
- `combine()`: aggregate groups of targets into other targets. Called inside `target()`.
- `file_in()`: declare an input file dependency.
- `file_out()`: declare an output file to be produced when the target is built.
- `knitr_in()`: declare a knitr file dependency such as an R Markdown (*.Rmd) or R LaTeX (*.Rnw) file.
- `ignore()`: force drake to entirely ignore a piece of code: do not track it for changes and do not analyze it for dependencies.
- `no_deps()`: tell drake to not track the dependencies of a piece of code. drake still tracks the code itself for changes.
- `drake_envir()`: get the environment where drake builds targets. Intended for advanced custom memory management.

See Also

`drake_plan()`, `make()`

Examples

```
# Use target() to create your own custom columns in a drake plan.
# See ?triggers for more on triggers.
drake_plan(
  website_data = target(
    download_data("www.your_url.com"),
    trigger = "always",
    custom_column = 5
  ),
  analysis = analyze(website_data)
)
models <- c("glm", "hierarchical")
plan <- drake_plan(
  data = target(
    get_data(x),
    transform = map(x = c("simulated", "survey"))
  ),
  analysis = target(
    analyze_data(data, model),
    transform = cross(data, model = !!models, .id = c(x, model))
  ),
  summary = target(
    summarize_analysis(analysis),
    transform = map(analysis, .id = c(x, model))
  ),
  results = target(
    bind_rows(summary),
    transform = combine(summary, .by = data)
  )
)
plan
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}
```

text_drake_graph

Show a workflow graph as text in your terminal window. **Stable**

Description

This is a low-tech version of `vis_drake_graph()` and friends. It is designed for when you do not have access to the usual graphics devices for viewing visuals in an interactive R session: for example, if you are logged into a remote machine with SSH and you do not have access to X Window support.

Usage

```
text_drake_graph(config, from = NULL, mode = c("out", "in", "all"),
  order = NULL, subset = NULL, targets_only = FALSE,
  make_imports = TRUE, from_scratch = FALSE, group = NULL,
  clusters = NULL, show_output_files = TRUE, nchar = 1L,
  print = TRUE)
```

Arguments

config	A drake_config() configuration list. You can get one as a return value from make() as well.
from	Optional collection of target/import names. If from is nonempty, the graph will restrict itself to a neighborhood of from. Control the neighborhood with mode and order.
mode	Which direction to branch out in the graph to create a neighborhood around from. Use "in" to go upstream, "out" to go downstream, and "all" to go both ways and disregard edge direction altogether.
order	How far to branch out to create a neighborhood around from. Defaults to as far as possible. If a target is in the neighborhood, then so are all of its custom file_out() files if show_output_files is TRUE. That means the actual graph order may be slightly greater than you might expect, but this ensures consistency between show_output_files = TRUE and show_output_files = FALSE.
subset	Optional character vector. Subset of targets/imports to display in the graph. Applied after from, mode, and order. Be advised: edges are only kept for adjacent nodes in subset. If you do not select all the intermediate nodes, edges will drop from the graph.
targets_only	Logical, whether to skip the imports and only include the targets in the workflow plan.
make_imports	Logical, whether to make the imports first. Set to FALSE to increase speed and risk using obsolete information.
from_scratch	Logical, whether to assume all the targets will be made from scratch on the next make() . Makes all targets outdated, but keeps information about build progress in previous make() s.
group	Optional character scalar, name of the column used to group nodes into columns. All the columns names of your original drake plan are choices. The other choices (such as "status") are column names in the nodes. To group nodes into clusters in the graph, you must also supply the clusters argument.
clusters	Optional character vector of values to cluster on. These values must be elements of the column of the nodes data frame that you specify in the group argument to drake_graph_info() .
show_output_files	Logical, whether to include file_out() files in the graph.
nchar	For each node, maximum number of characters of the node label to show. Can be 0, in which case each node is a colored box instead of a node label. Caution: nchar > 0 will mess with the layout.
print	Logical. If TRUE, the graph will print to the console via message() . If FALSE, nothing is printed. However, you still have the visualization because text_drake_graph() and render_text_drake_graph() still invisibly return a character string that you can print yourself with message() .

Value

A visNetwork graph.

See Also

[render_text_drake_graph\(\)](#), [vis_drake_graph\(\)](#), [sankey_drake_graph\(\)](#), [drake_ggraph\(\)](#)

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    config <- drake_config(my_plan)
    # Plot the network graph representation of the workflow.
    pkg <- requireNamespace("txtplot", quietly = TRUE) &&
      requireNamespace("visNetwork", quietly = TRUE)
    if (pkg) {
      text_drake_graph(config)
      make(my_plan) # Run the project, build the targets.
      text_drake_graph(config) # The black nodes from before are now green.
    }
  }
})

## End(Not run)
```

tracked

*List the targets and imports that are reproducibly tracked. **Stable***

Description

List all the layout in your project's dependency network.

Usage

```
tracked(config)
```

Arguments

config An output list from `drake_config()`.

Value

A character vector with the names of reproducibly-tracked targets.

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Load the canonical example for drake.
    # List all the targets/imports that are reproducibly tracked.
    config <- drake_config(my_plan)
    tracked(config)
  }
})

## End(Not run)
```

transform_plan	<i>Transform a plan</i> Maturing
----------------	---

Description

Evaluate the `map()`, `cross()`, `split()` and `combine()` operations in the `transform` column of a drake plan.

Usage

```
transform_plan(plan, envir = parent.frame(), trace = FALSE,
              max_expand = NULL, tidy_eval = TRUE)
```

Arguments

<code>plan</code>	A drake plan with a <code>transform</code> column
<code>envir</code>	Environment for tidy evaluation.
<code>trace</code>	Logical, whether to add columns to show what happens during target transformations.
<code>max_expand</code>	Positive integer, optional upper bound on the lengths of grouping variables for <code>map()</code> and <code>cross()</code> . Comes in handy when you have a massive number of targets and you want to test on a miniature version of your workflow before you scale up to production.
<code>tidy_eval</code>	Logical, whether to use tidy evaluation (e.g. unquoting/!!) when resolving commands. Tidy evaluation in transformations is always turned on regardless of the value you supply to this argument.

Details

<https://ropenscilabs.github.io/drake-manual/plans.html#large-plans> # nolint

See Also

`drake_plan`, `map`, `split`, `cross`, `combine`

Examples

```
plan1 <- drake_plan(
  y = target(
    f(x),
    transform = map(x = c(1, 2))
  ),
  transform = FALSE
)
plan2 <- drake_plan(
  z = target(
    g(y),
    transform = map(y, .id = x)
  ),
  transform = FALSE
)
```

```

plan <- bind_plans(plan1, plan2)
transform_plan(plan)
models <- c("glm", "hierarchical")
plan <- drake_plan(
  data = target(
    get_data(x),
    transform = map(x = c("simulated", "survey"))
  ),
  analysis = target(
    analyze_data(data, model),
    transform = cross(data, model = !!models, .id = c(x, model))
  ),
  summary = target(
    summarize_analysis(analysis),
    transform = map(analysis, .id = c(x, model))
  ),
  results = target(
    bind_rows(summary),
    transform = combine(summary, .by = data)
  )
)
plan
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}
# Tags:
drake_plan(
  x = target(
    command,
    transform = map(y = c(1, 2), .tag_in = from, .tag_out = c(to, out))
  ),
  trace = TRUE
)
plan <- drake_plan(
  survey = target(
    survey_data(x),
    transform = map(x = c(1, 2), .tag_in = source, .tag_out = dataset)
  ),
  download = target(
    download_data(),
    transform = map(y = c(5, 6), .tag_in = source, .tag_out = dataset)
  ),
  analysis = target(
    analyze(dataset),
    transform = map(dataset)
  ),
  results = target(
    bind_rows(analysis),
    transform = combine(analysis, .by = source)
  )
)
plan
if (requireNamespace("styler", quietly = TRUE)) {
  print(drake_plan_source(plan))
}

```

trigger

Customize the decision rules for rebuilding targets **Stable****Description**

Use this function inside a target's command in your `drake_plan()` or the `trigger` argument to `make()` or `drake_config()`. For details, see the chapter on triggers in the user manual: <https://ropenscilabs.github.io/drake-manual>

Usage

```
trigger(command = TRUE, depend = TRUE, file = TRUE, seed = TRUE,
        condition = FALSE, change = NULL, mode = c("whitelist",
        "blacklist", "condition"))
```

Arguments

command	Logical, whether to rebuild the target if the <code>drake_plan()</code> command changes.
depend	Logical, whether to rebuild if a non-file dependency changes.
file	Logical, whether to rebuild the target if a <code>file_in()/file_out()/knitr_in()</code> file changes.
seed	Logical, whether to rebuild the target if the seed changes. Only makes a difference if you set a custom seed column in your <code>drake_plan()</code> at some point in your workflow.
condition	R code (expression or language object) that returns a logical. The target will rebuild if the code evaluates to TRUE.
change	R code (expression or language object) that returns any value. The target will rebuild if that value is different from last time or not already cached.
mode	A character scalar equal to "whitelist" (default) or "blacklist" or "condition". With the mode argument, you can choose how the condition trigger factors into the decision to build or skip the target. Here are the options. <ul style="list-style-type: none"> • "whitelist" (default): we <i>rebuild</i> the target whenever condition evaluates to TRUE. Otherwise, we defer to the other triggers. This behavior is the same as the decision rule described in the "Details" section of this help file. • "blacklist": we <i>skip</i> the target whenever condition evaluates to FALSE. Otherwise, we defer to the other triggers. • "condition": here, the condition trigger is the only decider, and we ignore all the other triggers. We <i>rebuild</i> target whenever condition evaluates to TRUE and <i>skip</i> it whenever condition evaluates to FALSE.

Details

A target always builds if it has not been built before. Triggers allow you to customize the conditions under which a pre-existing target *rebuilds*. By default, the target will rebuild if and only if:

- Any of command, depend, or file is TRUE, or
- condition evaluates to TRUE, or

- `change` evaluates to a value different from last time. The above steps correspond to the "whitelist" decision rule. You can select other decision rules with the `mode` argument described in this help file. On another note, there may be a slight efficiency loss if you set complex triggers for `change` and/or `condition` because drake needs to load any required dependencies into memory before evaluating these triggers.

Value

A list of trigger specification details that drake processes internally when it comes time to decide whether to build the target.

See Also

[drake_plan\(\)](#), [make\(\)](#)

Examples

```
# A trigger is just a set of decision rules
# to decide whether to build a target.
trigger()
# This trigger will build a target on Tuesdays
# and when the value of an online dataset changes.
trigger(condition = today() == "Tuesday", change = get_online_dataset())
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    # You can use a global trigger argument:
    # for example, to always run everything.
    make(my_plan, trigger = trigger(condition = TRUE))
    make(my_plan, trigger = trigger(condition = TRUE))
    # You can also define specific triggers for each target.
    plan <- drake_plan(
      x = sample.int(15),
      y = target(
        command = x + 1,
        trigger = trigger(depend = FALSE)
      )
    )
    # Now, when x changes, y will not.
    make(plan)
    make(plan)
    plan$command[1] <- "sample.int(16)" # change x
    make(plan)
  }
})

## End(Not run)
```

Description

Add top-level R script files to use drake in your data analysis project. For details, read <https://ropenscilabs.github.io/drake-manual/projects.html>

Usage

```
use_drake(open = interactive())
```

Arguments

open Logical, whether to open make.R for editing.

Details

Files written:

1. make.R: a suggested master R script for batch mode.
2. _drake.R: a configuration R script for the `r_*`() functions documented at # nolint <https://ropenscilabs.github.io/drake-manual/projects.html#safer-interactivity>. # nolint Remarks:
 - There is nothing magical about the name, make.R. You can call it whatever you want.
 - Other supporting scripts, such as R/packages.R, R/functions.R, and R/plan.R, are not included.
 - You can find examples at <https://github.com/wlandau/drake-examples> and download examples with `drake_example()` (e.g. `drake_example("main")`).

Examples

```
## Not run:
# use_drake(open = FALSE) # nolint

## End(Not run)
```

vis_drake_graph	<i>Show an interactive visual network representation of your drake project.</i> Maturing
-----------------	---

Description

To save time for repeated plotting, this function is divided into `drake_graph_info()` and `render_drake_graph()`.

Usage

```
vis_drake_graph(config, file = character(0), selfcontained = FALSE,
  build_times = "build", digits = 3, targets_only = FALSE,
  font_size = 20, layout = NULL, main = NULL, direction = NULL,
  hover = FALSE, navigationButtons = TRUE, from = NULL,
  mode = c("out", "in", "all"), order = NULL, subset = NULL,
  ncol_legend = 1, full_legend = FALSE, make_imports = TRUE,
  from_scratch = FALSE, group = NULL, clusters = NULL,
  show_output_files = TRUE, collapse = TRUE, ...)
```

Arguments

config	A drake_config() configuration list. You can get one as a return value from make() as well.
file	Name of a file to save the graph. If NULL or character(0), no file is saved and the graph is rendered and displayed within R. If the file ends in a .png, .jpg, .jpeg, or .pdf extension, then a static image will be saved. In this case, the webshot package and PhantomJS are required: <code>install.packages("webshot"); webshot::install_phantomjs()</code> . If the file does not end in a .png, .jpg, .jpeg, or .pdf extension, an HTML file will be saved, and you can open the interactive graph using a web browser.
selfcontained	Logical, whether to save the file as a self-contained HTML file (with external resources base64 encoded) or a file with external resources placed in an adjacent directory. If TRUE, pandoc is required. The selfcontained argument only applies to HTML files. In other words, if file is a PNG, PDF, or JPEG file, for instance, the point is moot.
build_times	Character string or logical. If character, the choices are 1. "build": runtime of the command plus the time it take to store the target or import. 2. "command": just the runtime of the command. 3. "none": no build times. If logical, build_times selects whether to show the times from 'build_times(..., type = "build")' or use no build times at all. See build_times() for details.
digits	Number of digits for rounding the build times
targets_only	Logical, whether to skip the imports and only include the targets in the workflow plan.
font_size	Numeric, font size of the node labels in the graph
layout	Deprecated.
main	Character string, title of the graph.
direction	Deprecated.
hover	Logical, whether to show text (file contents, commands, etc.) when you hover your cursor over a node.
navigationButtons	Logical, whether to add navigation buttons with <code>visNetwork::visInteraction(navigationButtons = TRUE)</code>
from	Optional collection of target/import names. If from is nonempty, the graph will restrict itself to a neighborhood of from. Control the neighborhood with mode and order.
mode	Which direction to branch out in the graph to create a neighborhood around from. Use "in" to go upstream, "out" to go downstream, and "all" to go both ways and disregard edge direction altogether.
order	How far to branch out to create a neighborhood around from. Defaults to as far as possible. If a target is in the neighborhood, then so are all of its custom file_out() files if show_output_files is TRUE. That means the actual graph order may be slightly greater than you might expect, but this ensures consistency between show_output_files = TRUE and show_output_files = FALSE.
subset	Optional character vector. Subset of targets/imports to display in the graph. Applied after from, mode, and order. Be advised: edges are only kept for adjacent nodes in subset. If you do not select all the intermediate nodes, edges will drop from the graph.

ncol_legend	Number of columns in the legend nodes. To remove the legend entirely, set ncol_legend to NULL or 0.
full_legend	Logical. If TRUE, all the node types are printed in the legend. If FALSE, only the node types used are printed in the legend.
make_imports	Logical, whether to make the imports first. Set to FALSE to increase speed and risk using obsolete information.
from_scratch	Logical, whether to assume all the targets will be made from scratch on the next <code>make()</code> . Makes all targets outdated, but keeps information about build progress in previous <code>make()</code> s.
group	Optional character scalar, name of the column used to group nodes into columns. All the columns names of your original drake plan are choices. The other choices (such as "status") are column names in the nodes. To group nodes into clusters in the graph, you must also supply the clusters argument.
clusters	Optional character vector of values to cluster on. These values must be elements of the column of the nodes data frame that you specify in the group argument to <code>drake_graph_info()</code> .
show_output_files	Logical, whether to include <code>file_out()</code> files in the graph.
collapse	Logical, whether to allow nodes to collapse if you double click on them. Analogous to <code>visNetwork::visOptions(collapse = TRUE)</code> or <code>visNetwork::visOptions(collapse = TRUE)</code> .
...	Arguments passed to <code>visNetwork()</code> .

Value

A `visNetwork` graph.

See Also

`render_drake_graph()`, `sankey_drake_graph()`, `drake_ggraph()`, `text_drake_graph()`

Examples

```
## Not run:
isolate_example("Quarantine side effects.", {
  if (suppressWarnings(require("knitr"))){
    load_mtcars_example() # Get the code with drake_example("mtcars").
    config <- drake_config(my_plan)
    # Plot the network graph representation of the workflow.
    if (requireNamespace("visNetwork", quietly = TRUE)) {
      vis_drake_graph(config, width = '100%') # The width is passed to visNetwork
      make(my_plan) # Run the project, build the targets.
      vis_drake_graph(config) # The red nodes from before are now green.
      # Plot a subgraph of the workflow.
      vis_drake_graph(
        config,
        from = c("small", "reg2"),
        to = "summ_regression2_small"
      )
    }
  }
})
```

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

Index

`assign()`, [26](#), [67](#), [87](#)
`attachNamespace()`, [25](#), [66](#)

`bind_plans`, [4](#)
`build_times`, [5](#)
`build_times()`, [38](#), [40](#), [81–84](#), [101](#), [115](#)

`cached`, [6](#)
`cached()`, [23](#), [37](#), [85](#), [88](#), [97](#)
`clean`, [8](#)
`clean()`, [8](#), [10](#), [36](#), [97](#)
`clean_mtcars_example`, [9](#)
`clean_mtcars_example()`, [64](#)
`code_to_plan`, [10](#)
`code_to_plan()`, [80](#), [81](#)
`combine`, [11](#)
`combine()`, [12](#), [14](#), [33](#), [46](#), [47](#), [55](#), [57](#), [60](#), [62](#), [73](#), [77](#), [104–106](#)
`cross`, [13](#)
`cross()`, [12](#), [14](#), [33](#), [46](#), [47](#), [55](#), [56](#), [60](#), [62](#), [73](#), [77](#), [104–106](#)

`delayedAssign()`, [26](#), [68](#), [87](#)
`deps_code`, [15](#)
`deps_code()`, [16–18](#)
`deps_knitr`, [16](#)
`deps_knitr()`, [15](#), [18](#)
`deps_profile`, [16](#)
`deps_profile()`, [16](#)
`deps_target`, [17](#)
`deps_target()`, [15](#), [16](#), [87](#)
`diagnose`, [18](#)
`diagnose()`, [17](#), [37](#), [53](#), [85](#)
`drake` (drake-package), [4](#)
`drake-package`, [4](#)
`drake_build`, [19](#)
`drake_build()`, [32](#)
`drake_cache`, [21](#)
`drake_cache()`, [23](#), [25](#), [42](#), [66](#), [97](#)
`drake_cache_log`, [22](#)
`drake_config`, [24](#)
`drake_config()`, [17](#), [18](#), [21](#), [24](#), [25](#), [28](#), [30](#), [38](#), [40](#), [46](#), [66](#), [67](#), [70](#), [72](#), [75](#), [79](#), [82](#), [83](#), [87](#), [89–91](#), [99–101](#), [103](#), [106](#), [108](#), [109](#), [112](#), [115](#)
`drake_debug`, [31](#)
`drake_debug()`, [20](#)
`drake_envir`, [32](#)
`drake_envir()`, [12](#), [14](#), [29](#), [33](#), [47](#), [55](#), [57](#), [60](#), [62](#), [70](#), [74](#), [78](#), [104](#), [106](#)
`drake_example`, [34](#)
`drake_example()`, [35](#), [43](#), [44](#), [114](#)
`drake_examples`, [35](#)
`drake_examples()`, [34](#), [43](#), [44](#), [64](#)
`drake_gc`, [35](#)
`drake_gc()`, [9](#), [97](#)
`drake_get_session_info`, [37](#)
`drake_get_session_info()`, [85](#)
`drake_ggraph`, [38](#)
`drake_ggraph()`, [92–94](#), [96](#), [102](#), [108](#), [116](#)
`drake_graph_info`, [39](#)
`drake_graph_info()`, [91–95](#), [101](#), [114](#)
`drake_history`, [41](#)
`drake_history()`, [29](#), [36](#), [42](#), [71](#)
`drake_hpc_template_file`, [43](#)
`drake_hpc_template_file()`, [28](#), [44](#), [69](#)
`drake_hpc_template_files`, [44](#)
`drake_hpc_template_files()`, [44](#)
`drake_plan`, [45](#)
`drake_plan()`, [5](#), [7](#), [11](#), [12](#), [14](#), [19](#), [24](#), [30](#), [33](#), [37](#), [42](#), [45](#), [46](#), [49](#), [50](#), [55](#), [56](#), [59](#), [60](#), [62](#), [65](#), [66](#), [72](#), [73](#), [77](#), [79–81](#), [85](#), [87](#), [88](#), [91](#), [104–106](#), [112](#), [113](#)
`drake_plan_source`, [49](#)
`drake_plan_source()`, [49](#)
`drake_slice`, [50](#)

`expose_imports`, [51](#)

`failed`, [53](#)
`failed()`, [19](#), [98](#)
`file_in`, [54](#)
`file_in()`, [12](#), [14](#), [24](#), [33](#), [42](#), [47](#), [55](#), [57](#), [60](#), [62](#), [66](#), [74](#), [77](#), [78](#), [104](#), [106](#), [112](#)
`file_out`, [56](#)
`file_out()`, [9](#), [12](#), [14](#), [24](#), [33](#), [38–40](#), [42](#), [47](#), [55](#), [57](#), [60](#), [62](#), [66](#), [74](#), [77](#), [78](#), [102](#),

- [104, 106, 108, 112, 115, 116](#)
- `file_store`, [58](#)
- `file_store()`, [16](#)
- `find_cache`, [59](#)
- `from_plan()`, [33](#)
- `ignore`, [59](#)
- `ignore()`, [12, 14, 33, 47, 55, 57, 60, 62, 74, 77, 78, 87, 104, 106](#)
- `knitr_in`, [61](#)
- `knitr_in()`, [12, 14, 24, 33, 42, 47, 55, 57, 60, 62, 66, 74, 78, 104, 106, 112](#)
- `legend_nodes`, [63](#)
- `library()`, [18, 20, 25, 31, 66, 86, 99](#)
- `load_mtcars_example`, [64](#)
- `load_mtcars_example()`, [10](#)
- `loadadd (readd)`, [86](#)
- `loadadd()`, [6, 7, 16, 86, 87](#)
- `loadNamespace()`, [25, 66](#)
- `make`, [65](#)
- `make()`, [5, 7–9, 11, 17, 19, 21, 23, 24, 26, 27, 29, 30, 32, 34, 35, 37, 38, 40, 45, 46, 52–54, 59, 67–69, 71, 75, 76, 79–91, 97, 98, 100–102, 106, 108, 112, 113, 115, 116](#)
- `map`, [73](#)
- `map()`, [12, 14, 33, 46, 55, 56, 60, 62, 73, 77, 104–106](#)
- `missed`, [75](#)
- `missed()`, [79, 91](#)
- `new_cache`, [76](#)
- `new_cache()`, [5, 7, 8, 19, 21, 22, 25, 36, 37, 42, 53, 66, 85, 86, 89, 98](#)
- `no_deps`, [77](#)
- `no_deps()`, [12, 14, 33, 47, 55, 57, 60, 62, 74, 78, 104, 106](#)
- `outdated`, [78](#)
- `outdated()`, [30, 67, 71, 72, 75, 100](#)
- `plan_to_code`, [79](#)
- `plan_to_code()`, [10, 11, 80, 81](#)
- `plan_to_notebook`, [80](#)
- `plan_to_notebook()`, [10, 11, 80](#)
- `predict_runtime`, [81](#)
- `predict_runtime()`, [6, 25, 30, 66, 67, 84](#)
- `predict_workers`, [83](#)
- `predict_workers()`, [82](#)
- `progress`, [84](#)
- `progress()`, [19](#)
- `r_deps_target (r_make)`, [98](#)
- `r_drake_build (r_make)`, [98](#)
- `r_drake_ggraph (r_make)`, [98](#)
- `r_drake_graph_info (r_make)`, [98](#)
- `r_make`, [98](#)
- `r_make()`, [29, 71, 100](#)
- `r_missed (r_make)`, [98](#)
- `r_outdated (r_make)`, [98](#)
- `r_outdated()`, [71, 79, 91, 100](#)
- `r_predict_runtime (r_make)`, [98](#)
- `r_predict_workers (r_make)`, [98](#)
- `r_recoverable (r_make)`, [98](#)
- `r_recoverable()`, [29, 71, 91](#)
- `r_sankey_drake_graph (r_make)`, [98](#)
- `r_text_drake_graph (r_make)`, [98](#)
- `r_vis_drake_graph (r_make)`, [98](#)
- `r_vis_drake_graph()`, [99](#)
- `read_drake_seed`, [88](#)
- `readd`, [86](#)
- `readd()`, [6, 7, 16, 19, 37, 85–87](#)
- `recoverable`, [90](#)
- `recoverable()`, [29, 71](#)
- `remove()`, [7, 8, 85, 86](#)
- `render_drake_ggraph`, [91](#)
- `render_drake_ggraph()`, [39](#)
- `render_drake_graph`, [92](#)
- `render_drake_graph()`, [114, 116](#)
- `render_sankey_drake_graph`, [94](#)
- `render_sankey_drake_graph()`, [101, 102](#)
- `render_text_drake_graph`, [95](#)
- `render_text_drake_graph()`, [108](#)
- `require()`, [25, 66](#)
- `rescue_cache`, [96](#)
- `running`, [97](#)
- `running()`, [53](#)
- `sankey_drake_graph`, [101](#)
- `sankey_drake_graph()`, [39, 92–94, 96, 108, 116](#)
- `sessionInfo()`, [37](#)
- `shell_file()`, [44](#)
- `show_source`, [103](#)
- `show_source()`, [86](#)
- `split`, [103](#)
- `split()`, [12, 14, 33, 46, 55, 56, 60, 62, 73, 77, 104–106](#)
- `storr_rds()`, [76](#)
- `system.time()`, [6](#)
- `target`, [105](#)
- `target()`, [12, 14, 33, 46, 47, 55–57, 60, 62, 73, 77, 104, 106](#)
- `text_drake_graph`, [107](#)

`text_drake_graph()`, [39](#), [95](#), [96](#), [102](#), [116](#)
`tracked`, [109](#)
`transform_plan`, [110](#)
`trigger`, [112](#)
`trigger()`, [26](#), [67](#)

`use_drake`, [113](#)

`vis_drake_graph`, [114](#)
`vis_drake_graph()`, [30](#), [39](#), [41](#), [67](#), [72](#), [87](#),
 [92–94](#), [96](#), [99](#), [102](#), [107](#), [108](#)