

# Package: qtkit (via r-universe)

June 16, 2024

**Title** Quantitative Text Kit

**Version** 0.10.0

**Description** Support package for the textbook ``An Introduction to Quantitative Text Analysis for Linguists: Reproducible Research using R'' (Francom, 2024) <[doi:10.4324/9781003393764](https://doi.org/10.4324/9781003393764)> (available only after August 12, 2024). Includes functions to acquire, clean, and analyze text data as well as functions to document and share the results of text analysis. The package is designed to be used in conjunction with the book, but can also be used as a standalone package for text analysis.

**License** GPL (>= 3)

**URL** <https://qtaalr.github.io/qtkit/>, <https://github.com/qtaalr/qtkit>

**BugReports** <https://github.com/qtaalr/qtkit/issues>

**Depends** R (>= 3.6)

**Suggests** knitr, rmarkdown, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**Encoding** UTF-8

**Language** en-US

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.3.1

**VignetteBuilder** knitr

**Author** Jerid Francom [aut, cre, cph]  
(<<https://orcid.org/0000-0001-5972-6330>>)

**Maintainer** Jerid Francom <[francojc@wfu.edu](mailto:francojc@wfu.edu)>

**Repository** <https://francojc.r-universe.dev>

**RemoteUrl** <https://github.com/qtaalr/qtkit>

**RemoteRef** HEAD

**RemoteSha** 0f4495abab27e5c90631c9da41fe8f58c8885af8

## Contents

create_data_origin . . . . .	2
get_archive_data . . . . .	2

<b>Index</b>	<b>4</b>
--------------	----------

---

create_data_origin	<i>Create data origin file</i>
--------------------	--------------------------------

---

### Description

Data frame with attributes about the data origin, written to a CSV file and optionally returned.

### Usage

```
create_data_origin(file_path, return = FALSE, force = FALSE)
```

### Arguments

file_path	File path where the data origin file should be saved.
return	Logical value indicating whether the data origin should be returned.
force	Logical value indicating whether to overwrite the file if it already exists.

### Value

A data frame containing the data origin information.

### Examples

```
tmp_file <- tempfile(fileext = ".csv")
create_data_origin(tmp_file)
```

---

get_archive_data	<i>Download an archive file and unarchive its contents</i>
------------------	--

---

### Description

Possible file types include .zip, .gz, .tar, and .tgz

### Usage

```
get_archive_data(url, target_dir, force = FALSE, confirmed = FALSE)
```

**Arguments**

<code>url</code>	A character vector representing the full url to the compressed file
<code>target_dir</code>	The directory where the archive file should be downloaded
<code>force</code>	An optional argument which forcefully overwrites existing data
<code>confirmed</code>	If TRUE, the user has confirmed that they have permission to use the data. If FALSE, the function will prompt the user to confirm permission. Setting this to TRUE is useful for reproducible workflows.

**Value**

NULL, the archive file is unarchived in the target directory

**Examples**

```
test_dir <- file.path(tempdir(), "test")
url <- "https://raw.githubusercontent.com/qtalr/qtkit/main/inst/extdata/test_data.zip"
get_archive_data(
  url = url,
  target_dir = test_dir,
  confirmed = TRUE
)
```

# Index

`create_data_origin`, [2](#)

`get_archive_data`, [2](#)