

Package ‘laOpenData’

October 1, 2026

Type Package

Title Convenient Access to Los Angeles Open Data API Endpoints

Version 0.1.1

Description Provides simple, reproducible access to datasets from the Los Angeles Open Data portal <<https://data.lacity.org/>>. Functions return results as tidy tibbles and support optional filtering, sorting, and row limits via the Socrata API.

License MIT + file LICENSE

Encoding UTF-8

Imports dplyr, tibble, jsonlite, httr, janitor, rlang

Suggests curl, knitr, ggplot2, rmarkdown, testthat (>= 3.0.0), tidyr, vcr (>= 0.6.0), webmockr

URL <https://nyc-open-data-lab.github.io/laOpenData/>,
<https://github.com/nyc-open-data-lab/laOpenData>

BugReports <https://github.com/nyc-open-data-lab/laOpenData/issues>

VignetteBuilder knitr

Config/testthat/edition 3

Depends R (>= 4.1.0)

Config/roxygen2/version 8.1.0

NeedsCompilation no

Author Christian Martinez [aut, cre] (GitHub: martinezc1, ORCID:
<<https://orcid.org/0009-0005-6026-6454>>)

Maintainer Christian Martinez <c.martinez0@outlook.com>

Repository CRAN

Date/Publication 2026-10-01 19:20:02 UTC

Contents

la_any_dataset	2
la_list_datasets	3
la_pull_dataset	3

Index**5**

la_any_dataset	<i>Load Any Los Angeles Open Data Dataset</i>
----------------	---

Description

Downloads any Los Angeles Open Data dataset given its Socrata JSON endpoint.

Usage

```
la_any_dataset(
  json_link,
  limit = 10000,
  timeout_sec = 30,
  clean_names = TRUE,
  coerce_types = TRUE
)
```

Arguments

json_link	A Socrata dataset JSON endpoint URL (e.g., "https://data.lacity.org/resource/6rrh-rzua.json.json").
limit	Number of rows to retrieve (default = 10,000).
timeout_sec	Request timeout in seconds (default = 30).
clean_names	Logical; if TRUE, convert column names to snake_case (default = TRUE).
coerce_types	Logical; if TRUE, attempt light type coercion (default = TRUE).

Value

A tibble containing the requested dataset.

Examples

```
# Examples that hit the live Los Angeles Open Data API are guarded so CRAN checks
# do not fail when the network is unavailable or slow.
if (interactive() && curl::has_internet()) {
  endpoint <- "https://data.lacity.org/resource/6rrh-rzua.json.json"
  out <- try(la_any_dataset(endpoint, limit = 3), silent = TRUE)
  if (!inherits(out, "try-error")) {
    head(out)
  }
}
```

la_list_datasets	<i>List datasets available in laOpenData</i>
------------------	--

Description

Retrieves the current Open la catalog and returns datasets available for use with 'la_pull_dataset()'.

Usage

```
la_list_datasets()
```

Details

Keys are generated from dataset names using 'janitor::make_clean_names()'.

Value

A tibble of available datasets, including generated 'key', dataset 'id', and dataset 'name'.

Examples

```
if (interactive() && curl::has_internet()) {  
  la_list_datasets()  
}
```

la_pull_dataset	<i>Pull a Los Angeles Open Data dataset from the Los Angeles Open Data catalog</i>
-----------------	--

Description

Uses a dataset 'key' or 'id' from 'la_list_datasets()' to pull data from Los Angeles Open Data.

Usage

```
la_pull_dataset(  
  dataset,  
  limit = 10000,  
  filters = list(),  
  date = NULL,  
  from = NULL,  
  to = NULL,  
  date_field = NULL,  
  where = NULL,  
  order = NULL,  
  timeout_sec = 30,
```

```

    clean_names = TRUE,
    coerce_types = TRUE
  )

```

Arguments

<code>dataset</code>	A dataset key or ID from <code>'la_list_datasets()'</code> .
<code>limit</code>	Number of rows to retrieve (default = 10,000).
<code>filters</code>	Optional named list of filters. Supports vectors (translated to <code>IN()</code>).
<code>date</code>	Optional single date (matches all times that day) using <code>'date_field'</code> .
<code>from</code>	Optional start date (inclusive) using <code>'date_field'</code> .
<code>to</code>	Optional end date (exclusive) using <code>'date_field'</code> .
<code>date_field</code>	Optional date/datetime column to use with <code>'date'</code> , <code>'from'</code> , or <code>'to'</code> . Must be supplied when <code>'date'</code> , <code>'from'</code> , or <code>'to'</code> are used.
<code>where</code>	Optional raw SoQL WHERE clause. If <code>'date'</code> , <code>'from'</code> , or <code>'to'</code> are provided, their conditions are AND-ed with this.
<code>order</code>	Optional SoQL ORDER BY clause.
<code>timeout_sec</code>	Request timeout in seconds (default = 30).
<code>clean_names</code>	Logical; if TRUE, convert column names to snake_case (default = TRUE).
<code>coerce_types</code>	Logical; if TRUE, attempt light type coercion (default = TRUE).

Details

Dataset keys are generated from dataset titles using `'janitor::make_clean_names()'`. Because keys are derived from live catalog metadata, dataset IDs are the more stable option.

Value

A tibble.

Examples

```

if (interactive() && curl::has_internet()) {
  # Pull by key
  la_pull_dataset("current_employee_names_salaries_and_position_titles", limit = 3)

  # Pull by ID
  la_pull_dataset("xzkg-xp2w", limit = 3)

  # Filters
  la_pull_dataset("xzkg-xp2w", limit = 3, filters = list(salary_or_hourly = "HOURLY"))
}

```

Index

`la_any_dataset`, [2](#)
`la_list_datasets`, [3](#)
`la_pull_dataset`, [3](#)