

Package: plausibler (via r-universe)

September 18, 2024

Title Access Plausible Analytics API

Version 0.1.5

Description Access Plausible Analytics API.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

Imports curl, jsonlite, magrittr, purrr, tibble, tidyr

Suggests RSQLite, dplyr

URL <https://giocomai.github.io/plausibler/>

Repository <https://giocomai.r-universe.dev>

RemoteUrl <https://github.com/giocomai/plausibler>

RemoteRef HEAD

RemoteSha 78e3422ae9bfc7cdc222eb6354ba300840a79178

Contents

pa_create_handler	2
pa_get	2
pa_get_aggregate	3
pa_get_breakdown	4
pa_get_breakdown_by_date	5
pa_get_page_aggregate	6
pa_get_page_timeseries	7
pa_get_properties_by_date	7
pa_get_timeseries	9
pa_get_top_pages	10
pa_set	10

Index	12
--------------	-----------

pa_create_handler *Creates a curl handler object with API keys (used internally)*

Description

Creates a curl handler object with API keys (used internally)

Usage

```
pa_create_handler()
```

Value

A curl handler object

Examples

```
pa_set(
  base_url = "https://plausible.io/",
  site_id = "example.com",
  key = "actual_key_here"
)

pa_create_handler()
```

pa_get *Gets data from Plausible Analytics API*

Description

For details, consult the official documentation: <https://plausible.io/docs/stats-api>

Usage

```
pa_get(endpoint, parameters = NULL, filters = NULL, full_url = NULL)
```

Arguments

endpoint	Endpoint as described in the official documentation, e.g. <code>"/api/v1/stats/timeseries"</code> .
parameters	A named list with parameters (see example). If you want to run more complex queries you can leave this NULL, and include the whole call (Plausible instance URL, endpoint, and query) to the parameter <code>full_url</code> .
filters	Optional, defaults to NULL. If given, it must be given in the form <code>"visit:browser==Firefox;visit:country=="</code> or as a named vector (see examples). Use <code>;"</code> to separate multiple filtering criteria. For details, see the https://plausible.io/docs/stats-api#filtering for reference.
full_url	Defaults to NULL. If given, takes precedence over other parameters as well as settings. See examples, as well as examples in the official documentation.

Value

A data frame (a tibble) with results.

Examples

```
## Not run:
pa_get(
  endpoint = "/api/v1/stats/timeseries",
  parameters = list(period = "6mo")
)

# Same as above, but with full_url:
# N.B. You still need to set the key with pa_set(). Replace $SITE_ID with your domain
pa_get(full_url = "https://plausible.io/api/v1/stats/timeseries?site_id=$SITE_ID&period=6mo")

## End(Not run)
```

pa_get_aggregate	<i>Get aggregated stats for a given page</i>
------------------	----------------------------------------------

Description

Get aggregated stats for a given page

Usage

```
pa_get_aggregate(
  period = "30d",
  metrics = "visitors,pageviews,bounce_rate,visit_duration"
)
```

Arguments

period	Period to include in the analysis, defaults to "30d". See examples as well as the the official documentation for valid values: https://plausible.io/docs/stats-api#time-periods
metrics	Defaults to all available metrics. See documentation for more details: https://plausible.io/docs/stats-api#get-apiv1statsaggregate

Value

A data frame.

Examples

```
## Not run:
pa_get_aggregate(period = "6mo", page = "/")

## End(Not run)
```

pa_get_breakdown *Get a breakdown of stats by a given property*

Description

Get a breakdown of stats by a given property

Usage

```
pa_get_breakdown(
  period = "30d",
  property = "event:page",
  metrics = "visitors",
  limit = 100,
  filters = NULL
)
```

Arguments

period	Period to include in the analysis, defaults to "30d". See examples as well as the the official documentation for valid values: https://plausible.io/docs/stats-api#time-periods
property	Property to break down the stats by. For a full list of available properties, see the official documentation: https://plausible.io/docs/stats-api#properties
metrics	Default to "visitors". Can be set, for example, to c("visitors", "pageviews", "bounce_rate", "visit_duration"). For a full list of available metrics and their description, see the official documentation: https://plausible.io/docs/stats-api#metrics
limit	Limit the number of results. Maximum value is 1000. Defaults to 100. If you want to get more than 1000 results, you can make multiple requests and paginate the results by specifying the page parameter (e.g. make the same request with page=1, then page=2, etc)
filters	Optional, defaults to NULL. If given, it must be given in the form "visit:browser==Firefox;visit:country==" or as a named vector (see examples). Use ";" to separate multiple filtering criteria. For details, see the https://plausible.io/docs/stats-api#filtering for reference.

Value

A data frame.

Examples

```
## Not run:
pa_get_breakdown(period = "30d", property = "event:page")

## With filters, e.g. to see all referrers to the given url:
```

```

pa_get_breakdown(
  period = "30d",
  property = "visit:referrer",
  filters = list(`event:page` = "/berlin/")
)

## End(Not run)

```

pa_get_breakdown_by_date

Get breakdown by date and a single property

Description

This facilitates caching daily stats.

Usage

```

pa_get_breakdown_by_date(
  property = "event:page",
  metric = "visitors",
  start_date = Sys.Date() - 8,
  end_date = Sys.Date() - 1,
  limit = 1000,
  cache = TRUE,
  wait = 0.1
)

```

Arguments

start_date	Earliest day to include in the output. Defaults to 8 days ago in order to include the last full week of data. Date is expected in the format "YYYY-MM-DD", either as character or Date.
end_date	Most recent day to include in the output. Defaults to yesterday in order to include the last full week of data. Date is expected in the format "YYYY-MM-DD", either as character or Date.
limit	Defaults to 1000, to reduce the need for pagination, which is currently not supported.
cache	Defaults to TRUE. If TRUE, caches data in a local sqlite database, stored under a folder named as the website in the current working directory. The sqlite database is named after the chosen properties, hence there shouldn't be problems in caching data for different websites or different combinations of properties.
wait	Numeric, defaults to 0.1. As this function is likely to make a high number of requests to the API, a small pause is added between each request to reduce load on the servers. description

Details

However, the same data can mostly be retrieved more efficiently using the `timeseries` endpoint, setting the optional `interval` parameter to `date` (not yet integrated in the `pa_get_timeseries()` function).

Value

A data frame with three columns: `date`, `property`, and `metric`.

Examples

```
## Not run:
pa_get_breakdown_by_date()
pa_get_breakdown_by_date(property = "visit:source")

## End(Not run)
```

`pa_get_page_aggregate` *Get aggregated stats for a given page*

Description

Get aggregated stats for a given page

Usage

```
pa_get_page_aggregate(period = "30d", page)
```

Arguments

<code>period</code>	Period to include in the analysis, defaults to "30d". See examples as well as the the official documentation for valid values: https://plausible.io/docs/stats-api#time-periods
<code>limit</code>	Limit the number of results. Defaults to 100.

Value

A data frame.

Examples

```
## Not run:
pa_get_page_aggregate(period = "6mo", page = "/")

## End(Not run)
```

`pa_get_page_timeseries`*Get stats for a given page by time period*

Description

Get stats for a given page by time period

Usage

```
pa_get_page_timeseries(period = "30d", page)
```

Arguments

<code>period</code>	Period to include in the analysis, defaults to "30d". See examples as well as the the official documentation for valid values: https://plausible.io/docs/stats-api#time-periods
<code>limit</code>	Limit the number of results. Defaults to 100.

Value

A data frame.

Examples

```
## Not run:  
pa_get_page_timeseries(period = "6mo", page = "/")  
  
## End(Not run)
```

`pa_get_properties_by_date`*Get combined breakdown by two properties and date*

Description

This facilitates an operation that is not natively supported by the API.

Usage

```
pa_get_properties_by_date(
  property1 = "visit:source",
  property2 = "event:page",
  start_date = Sys.Date() - 8,
  end_date = Sys.Date() - 1,
  property1_to_exclude = character(),
  limit = 1000,
  cache = TRUE,
  wait = 0.1
)
```

Arguments

start_date	Earliest day to include in the output. Defaults to 8 days ago in order to include the last full week of data. Date is expected in the format "YYYY-MM-DD", either as character or Date.
end_date	Most recent day to include in the output. Defaults to yesterday in order to include the last full week of data. Date is expected in the format "YYYY-MM-DD", either as character or Date.
property1_to_exclude	Character vector. Useful to remove irrelevant iterations. For example, "Direct / None" should mostly be excluded when using visit:source or visit:referrer.
limit	Defaults to 1000, to reduce the need for pagination, which is currently not supported.
cache	Defaults to TRUE. If TRUE, caches data in a local sqlite database, stored under a folder named as the website in the current working directory. The sqlite database is named after the chosen properties, hence there shouldn't be problems in caching data for different websites or different combinations of properties.
wait	Numeric, defaults to 0.1. As this function is likely to make a high number of requests to the API, a small pause is added between each request to reduce load on the servers. description

Details

For example, this allows to retrieve visit:source for each event:page for each date. As this is not supported by the API, it achieves this result via repeated calls to the API: first it retrieves all visitors for property1 (e.g. for all visit:source) for a given date, then it iterates through each of them, and, relying on filters, queries one by one for each item resulting from this query (e.g. each visit:source found on a given date).

See <https://github.com/plausible/analytics/discussions/1254>

Value

A data frame with four columns: date, property1, property2, and visitors.

Examples

```
## Not run:  
pa_get_properties_by_date()  
pa_get_properties_by_date(property1 = "visit:referrer")  
  
## End(Not run)
```

pa_get_timeseries *Get time series*

Description

Get time series

Usage

```
pa_get_timeseries(period = "30d")
```

Arguments

period	Period to include in the analysis, defaults to "30d". See examples as well as the the official documentation for valid values: https://plausible.io/docs/stats-api#time-periods
--------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Value

A data frame.

Examples

```
## Not run:  
pa_get_timeseries(period = "6mo")  
pa_get_timeseries(period = "30d")  
pa_get_timeseries(period = "custom&date=2021-06-01,2021-06-07")  
  
## End(Not run)
```

pa_get_top_pages *Get most visited pages for a given period*

Description

Get most visited pages for a given period

Usage

```
pa_get_top_pages(period = "30d", limit = 100)
```

Arguments

period	Period to include in the analysis, defaults to "30d". See examples as well as the the official documentation for valid values: https://plausible.io/docs/stats-api#time-periods
limit	Limit the number of results. Defaults to 100.

Value

A data frame.

Examples

```
## Not run:
pa_get_top_pages(period = "6mo")

## End(Not run)
```

pa_set *Set API key for the current session*

Description

Set API key for the current session

Usage

```
pa_set(base_url = NULL, site_id = NULL, key = NULL)
```

Arguments

base_url	The base URL of the Plausible instance. Set to https://plausible.io/ - or to your own domain for self-hosted Plausible.
site_id	Corresponds to the domain of your website.
key	A character string used for authentication. Can be retrieved from the user settings in Plausible Analytics.

Value

Invisibly returns input as list.

Examples

```
pa_set(  
  base_url = "https://plausible.io/",  
  site_id = "example.com",  
  key = "actual_key_here"  
)
```

Index

pa_create_handler, [2](#)
pa_get, [2](#)
pa_get_aggregate, [3](#)
pa_get_breakdown, [4](#)
pa_get_breakdown_by_date, [5](#)
pa_get_page_aggregate, [6](#)
pa_get_page_timeseries, [7](#)
pa_get_properties_by_date, [7](#)
pa_get_timeseries, [9](#)
pa_get_top_pages, [10](#)
pa_set, [10](#)