Package 'DemografixeR'

October 12, 2022

Title Extrapolate Gender, Age and Nationality of a Name

Version 0.1.1

Description Connects to the <https://genderize.io/>, <https://agify.io/> and <https://nationalize.io/> APIs to estimate gender, age and nationality of a first name.

License MIT + file LICENSE

URL <https://matbmeijer.github.io/DemografixeR>,
<https://github.com/matbmeijer/DemografixeR>

BugReports <https://github.com/matbmeijer/DemografixeR/issues>

Encoding UTF-8

LazyData true

Imports httr, jsonlite, magrittr

RoxygenNote 7.1.0

Suggests covr, testthat, knitr, rmarkdown, dplyr

VignetteBuilder knitr

NeedsCompilation no

Author Matthias Brenninkmeijer [aut, cre]

Maintainer Matthias Brenninkmeijer <matthias@brennotten.net>

Repository CRAN

Date/Publication 2020-05-06 12:50:22 UTC

R topics documented:

---

agify   *'Agifies' a name*

Description

Function to estimate the age from a first name. Connects directly to the agify.io API sending a request with a name and parses the response to return the predicted age.

Usage

```
agify(
  name,
  country_id = NULL,
  simplify = TRUE,
  apikey = get_api_key(),
  meta = FALSE
)
```

Arguments

name
: Name/s to estimate the age. Can be a single character string or a character vector. Obligatory parameter.

country_id
: Responses will in a lot of cases be more accurate if the data is narrowed to a specific country. This optional parameter allows to specify a specific country. The parameter must be a country code following the common ISO 3166-1 alpha-2 country code convention. To see a list of the supported countries use the supported_countries function or visit the following link.

simplify
: Logical parameter, which defines if the result should be returned as a character vector or a data.frame with additional information. By default set to TRUE, which returns a vector.

apikey
: Optional parameter to pass the API key. The API is free for up to 1000 names/day. No sign up or API key needed. Yet, if more requests would be needed, visit the agify.io store and the obtained API key can be passed through this parameter. The API can also be saved one time through the save_api_key function, so it is not necessary to call again.

meta
: Logical parameter to define if API related information should be returned. By default set to FALSE. Returns information about:

    - The amount of names available in the current time window
    - The number of names left in the current time window
    - Seconds remaining until a new time window opens

Details

The function automatically handles pagination (max. 10 names per API request), missing values & duplicated values. If a name is duplicated for the same country_id (if given), it will handle the request as a a single name to save requests, returning the same value for the duplicated names. To see more details about the API documentation, visit the agify.io website.

Value

The estimated age in a single character vector form or a data.frame with additional information.

Warning

Please be aware about local privacy protection regulations such as GDPR when dealing with personal data.

Author(s)

Matthias Brenninkmeijer - https://github.com/matbmeijer

See Also

genderize, nationalize

Examples

```
agify(name=c("Ben", "Maria"))
```

DemografixeR   *DemografixeR: Estimate gender, age and nationality from a name*

Description

The DemografixeR is an API wrapper for the genderize.io, agify.io & nationalize.io APIs. Through the APIs it retrieves the estimated gender, estimated age and/or estimated nationality from a name. The package most important functions are the following:

- genderize
- agify
- nationalize

---

genderize	*'Genderizes' a name*

Description

Function to estimate the gender from a first name. Connects directly to the genderize.io API sending a request with a name and parses the response to return the predicted gender.

Usage

```
genderize(
  name,
  country_id = NULL,
  simplify = TRUE,
  apikey = get_api_key(),
  meta = FALSE
)
```

Arguments

name	Name/s to estimate the gender. Can be a single `character` string or a `character` vector. Obligatory parameter.
country_id	Responses will in a lot of cases be more accurate if the data is narrowed to a specific country. This optional parameter allows to specify a specific country. The parameter must be a country code following the common ISO 3166-1 alpha-2 country code convention. To see a list of the supported countries use the `supported_countries` function or visit the following link.
simplify	Defines if the result should be returned as a single vector or a `data.frame` with additional information. By default set to `TRUE`, which returns a vector. If set to `FALSE`, it will include additional information about:

- The probability indicating the certainty of the assigned gender. Basically the ratio of male to females.
- The count representing the number of data rows examined in order to calculate the response.

apikey	Optional parameter to pass the API key. The API is free for up to 1000 names/day. No sign up or API key needed. Yet, if more requests would be needed, visit the genderize.io store and the obtained API key can be passed through this parameter. The API can also be saved one time through the `save_api_key` function, so it is not necessary to call again.
meta	`Logical` parameter to define if API related information should be returned. By default set to `FALSE`. Returns information about:

- The amount of names available in the current time window
- The number of names left in the current time window
- Seconds remaining until a new time window opens

Details

The function automatically handles pagination (max. 10 names per API request), missing values & duplicated values. If a name is duplicated for the same country_id (if the parameter has been passed), it will handle the request as a a single name to save requests, returning the same value for the duplicated names. To see more details about the API documentation, visit the genderize.io website.

Value

The estimated age in a single `character` vector form or a `data.frame` with additional information.

Warning

Please be aware about local privacy protection regulations such as GDPR when dealing with personal data.

Author(s)

Matthias Brenninkmeijer - https://github.com/matbmeijer

See Also

nationalize, agify

Examples

```
genderize(name=c("Ben", "Maria"))
```

---

get_api_key   *Get previously saved API keys*

Description

Function to get the previously saved API key. It has the advantage that it is not necessary to explicitly publish the key in the users code. Just do it one time and you're set. To save an API, use the save_api_key function. To remove a previously saved key, use the remove_api_key function.

Usage

```
get_api_key()
```

Value

Returns the saved API key in the environment. If no API key has been saved, returns `NULL` value.

Warning

Please be careful when dealing with API keys and other secrets & tokens - keep them private and do not publish them.

Author(s)

Matthias Brenninkmeijer - https://github.com/matbmeijer

See Also

save_api_key, remove_api_key

Examples

```
## Not run:
get_api_key(type="genderize")
get_api_key(type="agify")
get_api_key(type="nationalize")

## End(Not run)
```

---

nationalize	*'Nationalizes' a name*

Description

Returns the estimated nationality from a given name.

Usage

```
nationalize(
  name,
  simplify = TRUE,
  sliced = TRUE,
  apikey = get_api_key(),
  meta = FALSE
)
```

Arguments

name	Name/s to estimate the nationality. Can be a single `character` string or a `character` vector. Obligatory parameter.
simplify	Defines if the result should be returned as a single vector or a `data.frame` with additional information. By default set to `TRUE`, which returns a vector.
sliced	Names can have multiple estimated nationalities ranked by probability. This `logical` parameter allows to "slice"/keep only the parameter with the highest probability to keep a single estimate for each name. By default set to `TRUE`.

apikey	Optional parameter to pass the API key. The API is free for up to 1000 names/day. No sign up or API key needed. Yet, if more requests would be needed, visit the nationalize.io store and the obtained API key can be passed through this parameter. The API can also be saved one time through the `save_api_key` function, so it is not necessary to call again.
meta	`Logical` parameter to define if API related information should be returned. By default set to `FALSE`. Returns information about:

- The amount of names available in the current time window
- The number of names left in the current time window
- Seconds remaining until a new time window opens

Details

The function automatically handles pagination (max. 10 names per API request), missing values & duplicated values. If a name is duplicated for the same `country_id` (if the parameter has been passed), it will handle the request as a a single name to save requests, returning the same value for the duplicated names. To see more details about the API documentation, visit the nationalize.io website.

Value

Returns the estimated nationality as the common ISO 3166-1 alpha-2 country code. The result can be in a single `character` vector form or a `data.frame` with additional information.

Warning

Please be aware about local privacy protection regulations such as GDPR when dealing with personal data.

Author(s)

Matthias Brenninkmeijer - https://github.com/matbmeijer

See Also

genderize, agify

Examples

```
nationalize(name=c("Ben", "Maria"))
```

---

remove_api_key *Removes saved API key*

Description

Removes saved API keys for the DemografixeR APIs (Genderize.io, Agify.io, Nationalize.io).

Usage

```
remove_api_key(verbose = TRUE)
```

Arguments

verbose   Logical parameter to define if a verbose message should be printed. By default
               set to TRUE.

Value

Does not return any object.

Warning

Please be careful when dealing with API keys and other secrets & tokens - keep them private and
do not publish them.

Author(s)

Matthias Brenninkmeijer - https://github.com/matbmeijer

See Also

save_api_key, get_api_key

Examples

```
## Not run:
remove_api_key()

## End(Not run)
```

---

save_api_key *Saves the API key for future functions*

Description

Saves the all in one genderize.io, agify.io & nationalize.io API key in the users environment. It has the advantage that it is not necessary to explicitly publish the key in the users code. Just do it one time and you're set. To update the key just save again and it will overwrite the old key. To explicitly print the key, use the get_api_key function. To remove the key use the remove_api_key function.

Usage

```
save_api_key(key)
```

Arguments

key API key obtained from the specific website. Visit the one of the following web-
sites to obtain an API key:

- genderize.io
- agify.io
- nationalize.io

Value

Does save the key in the environment.

Warning

Please be careful when dealing with API keys and other secrets & tokens - keep them private and do not publish them.

Author(s)

Matthias Brenninkmeijer - https://github.com/matbmeijer

See Also

get_api_key, remove_api_key

Examples

```
## Not run:
save_api_key(key="__YOUR_API_KEY__")

## End(Not run)
```

---

supported_countries   *Retrieve the supported countries for each API*

Description

Scrapes the API websites to retrieve an updated list of supported countries for each API.

Usage

```
supported_countries(type)
```

Arguments

type   Obligatory parameter to define the API from which to obtain the supported coun-
                  tries. Must be one of the following character strings:

                        • genderize - Available countries for the Genderize.io API.
                        • agify - Available countries for the Agify.io API.
                        • nationalize - Available countries for the Nationalize.io API.

Value

Returns a data.frame with the supported ISO 3166-1 alpha-2 country code, country name the
number of items (names) for each country.

Warning

Please be conscious that this function directly scrapes the website, do not overuse it as it might
overwhelm the server.

Examples

```
## Not run:
supported_countries(type="genderize")

## End(Not run)
```

Index