eia_backfill  

Pull a Large Number of Observations with a Sequential Query

Description
This function allows users to overcome the API’s observation limit per query by breaking down the query into smaller sequential sub-queries and appending back the results. The main use case of this function is for backfilling hourly series.

Usage

eia_backfill(start, end, offset, api_key, api_path, facets)

Arguments

start  
defines the start time of the series, should use a POSIXct class for hourly series or Date format for non-hourly series (daily, monthly, etc.)

defines the end time of the series, should use a POSIXct class for hourly series or Date format for non-hourly series (daily, monthly, etc.)

offset  
An integer, defines the number of observations limitation per query. In line with the API limitation of up to 5000 observations per query, the offset argument’s upper limit is 5000 observations.

api_key  
A string, EIA API key, see https://www.eia.gov/opendata/ for registration to the API service

api_path  
A string, the API path to follow the API endpoint https://api.eia.gov/v2/. The path can be found on the EIA API dashboard, for more details see https://www.eia.gov/opendata/browser/

facets  
A list, optional, set the filtering argument (defined as ‘facets’ on the API header), following the structure of list(facet_name_1 = value_1, facet_name_2 = value_2)

Details
The function use start, end, and offset arguments to define a sequence of queries.

Value

A time series

Examples

## Not run:
start <- as.POSIXlt("2018-06-19T00", tz = "UTC")
end <- lubridate::floor_date(Sys.time()- lubridate::days(2), unit = "day")
attr(end, "tzone") <- "UTC"
offset <- 2000
api_key <- Sys.getenv("eia_key")
api_path <- "electricity/rto/region-sub-ba-data/data/"
facets = list(parent = "NYIS", subba = "ZONA")

df <- eia_backfill(start = start, end = end, offset = offset, api_key = api_key, api_path = api_path, facets = facets)

at_y <- pretty(df$value)[c(2, 4, 6)]
at_x <- seq.POSIXt(from = start, to = end, by = "2 years")

plot(df$time, df$value, col = "#1f77b4", type = "l", frame.plot = FALSE, axes = FALSE, panel.first = abline(h = at_y, col = "grey80"), main = "NY Independent System Operator (West) - Hourly Generation of Electricity", xlab = "Source: https://www.eia.gov/", ylab = "MegaWatt/Hours")

mtext(side =1, text = format(at_x, format = "%Y"), at = at_x, col = "grey20", line = 1, cex = 0.8)
mtext(side =2, text = format(at_y, scientific = FALSE), at = at_y, col = "grey20", line = 1, cex = 0.8)

## End(Not run)

eia_get

### Query the EIA API

**Description**

Function to query and extract data from the EIA API v2

**Usage**

eia_get(
  api_key,
  api_path,
  data = "value",
  facets = NULL,
  start = NULL,
  end = NULL,
  length = NULL,
offset = NULL,
frequency = NULL,
format = "data.frame"
)

Arguments

api_key A string, EIA API key, see https://www.eia.gov/opendata/ for registration to the API service

api_path A string, the API path to follow the API endpoint https://api.eia.gov/v2/. The path can be found on the EIA API dashboard, for more details see https://www.eia.gov/opendata/browser/

data A string, the metric type, by default uses 'value' (defined as 'data' on the API header)

facets A list, optional, set the filtering argument (defined as 'facets' on the API header), following the structure of list(facet_name_1 = value_1, facet_name_2 = value_2)

start A string, optional, set the starting date or time of the series using "YYYY-MM-DD" format for date and "YYYY-MM-DDTHH" format for hourly time series

end A string, optional, set the ending date or time of the series using "YYYY-MM-DD" format for date and "YYYY-MM-DDTHH" format for hourly time series

length An integer, optional, defines the length of the series, if set to NULL (default), will default to the API default value of 5000 observations per pull. The API enables a pull of up to 100K observations per call. If needed to pull more than the API limit per call, recommend to iterate the call with the use of the start, end and/or offset arguments

offset An integer, optional, set the number of observations to offset from the default starting point of the series. If set to NULL (default), will default to the API default value of 0

frequency A string, optional, define the API frequency argument (e.g., hourly, monthly, annual, etc.). If set to NULL (default), will default to the API default value

format A string, defines the output of the return object to either "data.frame" (default) or "data.table"

Value
data.table/data.frame object

Examples

## Not run:
# Required an EIA API key to send a query
api_key <- "YOUR_API_KEY"

df <- eia_get(
  api_key = api_key,
  api_path = "electricity/rto/fuel-type-data/data/",
  data = "value"
)
## Pull Metadata from EIA API

### Description
Get data descriptions and metadata from the EIA API

### Usage
```r
eia_metadata(api_path = NULL, api_key)
```

### Arguments
- **api_path**: A string, the API category/route path following the API endpoint (i.e., 'https://api.eia.gov/v2/') If set to NULL (default) or as empty string "" it returns the main categories available on the API. The path can be found on the EIA API dashboard, for more details see https://www.eia.gov/opendata/browser/
- **api_key**: A string, EIA API key, see https://www.eia.gov/opendata/ for registration to the API service

### Details
The function enables to explore the different data categories and available routes inline with the API dashboard (https://www.eia.gov/opendata/browser/)

### Value
a list object with the series description and metadata

### Examples
```r
## Not run:
electricity_metadata <- eia_metadata(api_key = Sys.getenv("eia_key"),
                      api_path = "electricity")

electricity_metadata$response$description
electricity_metadata$response$id
electricity_metadata$response$name
electricity_metadata$response(routes

## End(Not run)
```
Index

eia_backfill, 2
eia_get, 3
eia_metadata, 5