Package 'odataR'

May 5, 2016

Type Package

Title Access to OData information

Version 0.1.2

Author Han Oostdijk

Maintainer <han.oostdijk@xs4all.nl>

Description Access to OData information.

License MIT + file LICENSE

LazyData TRUE

RoxygenNote 5.0.1.9000

Imports XML,
curl,
dplyr,
magrittr

R topics documented:

---

odataR   *odataR: A package for accessing OData information.*

Description

The odataR package contains some functions to obtain information from databases with the OData interface.

Details

Located under the root of an OData database are some tables. To obtain the information one should specify the correct root and the name of the table. Because this package was created to access the information of the CBS (Statistics Netherlands) the default root is appropriate for this information. The name of a table (here called table_id) can be found in the documentation of the database. For the CBS you can use the catalog http://opendata.cbs.nl/dataportaal/portal.html to see the available tables and their metadata. It is also possible to get the catalog information in a data.frame by using the function odataR_list_tables.

Setting a different root can be done by using the function odataR_set_root and querying the root by executing odataR_get_root.

The other functions in the package are related to actually obtain information from the database. In the CBS database a table name (table_id) points to a set of sub tables that together provide the information. One 'main' sub table contains the topic data and dimensions in coded form indicating where the topic data relates to. The other sub tables convey the meaning of the coded dimensions. E.g. a topic field could be the number of unemployed workers and the dimensions could be gender, age-class, province and year.

The main function is odataR_get_table that extracts the information from the database and decodes the dimensions. The other functions are used by odataR_get_table but can also be used independently. E.g odataR_get_subtable can be used to get the information in encoded form (needs less memory) and with odataR_get_subtable(subtabs['TableInfos'],mt=NULL) one gets the table information in XML form (however the information in the catalog is better readable for a human.)

For examples see odataR_get_table

---

odataR_get_data   *Get (part of) an OData data structure*

Description

This is the only function in the package that actually acquires data from the database. This function returns a block of at most 10000 rows (restriction of the interface) in XML format. It is therefore also suitable in cases where the result is not a nice table (e.g. in the case of TypeInfo). The function odataR_get_table will call this function one or more times (depending on the size of the data) and convert the combined data to a table or property structure.

Usage

```
odataR_get_data(root = odataR_get_root_data(), table_id = NULL,
  query = NULL, save_file_name = NULL)
```

Arguments

`root`	Root of data structure
`table_id`	Identification of table
`query`	OData query to restrict data returned from structure, See [`odataR_get_table`](#) for examples of queries
`save_file_name`	Name of file to save the XML data in or NULL

Value

An `XMLInternalDocument` if successful otherwise NULL

Examples

```
## Not run:
doc = odataR_get_data(table_id='82935NED')
doc = odataR_get_data(root='http://opendata.cbs.nl/ODataFeed/OData',table_id='82935NED')

## End(Not run)
```

---

odataR_get_root	*Gets the root for OData data and catalog structure*

Description

Determines from which structure data and catalog information will be extracted. If the function odataR_set_root is not called yet this will be done first with the default `http://opendata.cbs.nl` (the root for the CBS Statistics Netherlands structures).

Usage

```
odataR_get_root()
```

See Also

[`odataR_set_root`](#)

Examples

```
odataR_get_root()
```

---

odataR_get_root_catalog

Gets the root for OData catalog structure

Description

Determines from which structure the catalog information will be extracted. Derived from the common OData data and catalog root: paste0(odataR_get_root(), '/ODataCatalog/Tables')

Usage

```
odataR_get_root_catalog()
```

See Also

[odataR_get_root](#) and [odataR_get_root_data](#)

---

odataR_get_root_data   *Gets the root for OData data structure*

Description

Determines from which structure the data will be extracted. Derived from the common OData data and catalog root: paste0(odataR_get_root(), '/ODataFeed/OData')

Usage

```
odataR_get_root_data()
```

See Also

[odataR_get_root](#) and [odataR_get_root_catalog](#)

---

odataR_get_subtable   *Convert a sub table (contents or properties) to a data.frame*

Description

Convert a sub table (contents or properties) to a data.frame. The function calls [odataR_get_data](#) one or more times (the latter if the server returns more than 10000 rows) to actually retrieve the XML data. With parameter mt one indicates that the requests relates to a 'normal' sub table or relates to the properties sub table. In these cases the data is converted to a data.frame . By specifying mt=NULL the contents is not converted to a data.frame but returned as XML.

Usage

```
odataR_get_subtable(dsn, mt = "table", query = NULL, save_XML = NULL)
```

Arguments

dsn	Named character vector with full name of subtable
mt	Indicates destination of the XML data. Character `'table'` (default) or `'prop'` or NULL
query	OData query to restrict data returned from structure, See `odataR_get_table` for examples of queries
save_XML	Name of file to save the XML data in or NULL or default name

Value

A data.frame with the contents of the sub table (`mt='table'`) or the properties of the table (`mt='prop'`) or the XML contents of the sub table (`mt=NULL`)

Examples

```
## Not run:
subtabs = odataR_get_subtables(table_id="82935NED")

## End(Not run)
```

---

odataR_get_subtables   *Get information about sub tables in OData data structure*

Description

A logical table consists of a some related sub tables. This function returns a named character vector where the names give for each sub table the type and the contents the url of the sub table. This information can be used by the function `odataR_get_subtable` to retrieve the contents of a sub table

Usage

```
odataR_get_subtables(root = odataR_get_root_data(), table_id = NULL)
```

Arguments

root	Root of data structure
table_id	Identification of table

Value

A named character vector where the names give for each sub table the type and the contents the url of the sub table

Examples

```
## Not run:
subtabs = odataR_get_subtables(table_id='82935NED')
subtabs = odataR_get_subtables(root='http://opendata.cbs.nl/ODataFeed/OData',table_id='82935NED')

## End(Not run)
```

---

odataR_get_table   *Get table in decoded form*

Description

The main function is odataR_get_table that extracts the information from the database and de-codes the dimensions. It uses the other functions in the package behind the screens. E.g. when executing df = odataR_get_table(table_id="82935NED") the following is executed:

- get names of sub tables:
  subtabs = odataR_get_subtables(table_id="82935NED")

- retrieve coded data:
  df   = odataR_get_subtable(subtabs['TypedDataSet'])

- retrieve properties:
  props   = odataR_get_subtable(subtabs['DataProperties'],mt='prop')

- decode dimensions:
  df   = couple_data(df, ...)

Usage

```
odataR_get_table(root = odataR_get_root_data(), table_id = NULL,
  query = NULL, typed = T, keepcode = c())
```

Arguments

root	Root of data structure
table_id	Identification of table
query	OData query to restrict data returned from structure
typed	Boolean indicating 'TypedDataSet' when T or 'UntypedDataSet' when F'
keepcode	Character string with dimension(s) for which the coded values are kept

Value

A data.frame when successful otherwise NULL

Remark

See http://docs.oasis-open.org/odata/odata/v4.0/errata02/os/complete/part1-protocol/odata-v4.0-errata02-os-part1-protocol-complete.html for details about the query possibilities

$format=atom is default so can be omitted, but specifying $format=json will not work in this context

Examples

```
## Not run:
df      = odataR_get_table(table_id="82935NED")
df      = odataR_get_table(table_id="82935NED",keepcode = "RegioS")
df      = odataR_get_table(table_id="82935NED",keepcode = c("RegioS","Perioden"))
df      = odataR_get_table(table_id="82935NED",keepcode = "RegioS",
      query  = "?$format=atom&$filter=startswith(RegioS,'NL01')" )
df      = odataR_get_table(table_id="82935NED",
  query  = paste0("?$filter=startswith(RegioS,'NL01')",
                  "&$select=RegioS,Perioden,TotaleInvesteringen_1") )
df      = odataR_get_table(table_id="82935NED",
  query  = paste0("?$filter=startswith(RegioS,'NL01')",
                  "&$select=RegioS,Perioden,TotaleInvesteringen_1",
                  "&$skip=2&$top=3") )

## End(Not run)
```

---

odataR_list_tables   *List the tables in an OData structure*

Description

List the tables in an OData structure

Usage

```
odataR_list_tables(root = odataR_get_root_catalog(), query = NULL)
```

Arguments

root   Root of data structure

query   OData query to restrict data returned from structure, See odataR_get_table
                for examples of queries

Value

A data.frame when successful otherwise NULL

Examples

```
## Not run:
odataR_set_root("http://dataderden.cbs.nl")
df_cat = odataR_list_tables()
RIVM   = odataR_list_tables(query="?$filter=Catalog eq 'RIVM'")

## End(Not run)
```

| odataR_set_root | *Sets the root for OData data and catalog structure* |

Description

Determines from which structure data and catalog information will be extracted. If this function is not called the default `http://opendata.cbs.nl` (the root for the CBS Statistics Netherlands structures) will be used

Usage

```
odataR_set_root(root = NULL)
```

Arguments

root   NULL for the default or the url of the root structure otherwise

See Also

[odataR_get_root](#)

Examples

```
odataR_set_root()
odataR_set_root('http://opendata.cbs.nl')
```

Index