

Package ‘hmrc’

March 12, 2026

Title Download and Tidy HMRC Statistical Data

Version 0.3.0

Description Provides functions to download, parse, and tidy statistical data published by HM Revenue and Customs (HMRC) on GOV.UK. Covers monthly tax receipts (41 tax heads from 2016), VAT (from 1973), fuel duties (from 1990), tobacco duties (from 1991), annual Corporation Tax receipts, stamp duty, research and development tax credit statistics (from 2000), tax gap estimates, Income Tax liabilities by income range, and monthly property transaction counts. File URLs are resolved at run-time via the GOV.UK Content API <https://www.gov.uk/api/content>, so data is always current without hardcoded URLs. Files are cached locally between sessions.

License MIT + file LICENSE

Encoding UTF-8

Language en-GB

RoxygenNote 7.3.3

Imports cli (>= 3.6.0), httr2 (>= 1.0.0), readODS (>= 2.3.0), readxl (>= 1.4.0), tools

Suggests knitr, rmarkdown, testthat (>= 3.0.0), ggplot2, scales

Config/testthat/edition 3

VignetteBuilder knitr

URL <https://github.com/charlescoverdale/hmrc>

BugReports <https://github.com/charlescoverdale/hmrc/issues>

Depends R (>= 4.1.0)

LazyData true

NeedsCompilation no

Author Charles Coverdale [aut, cre]

Maintainer Charles Coverdale <charlesfcoverdale@gmail.com>

Repository CRAN

Date/Publication 2026-03-12 08:50:02 UTC

Contents

clear_cache	2
get_corporation_tax	3
get_fuel_duties	4
get_income_tax_stats	5
get_property_transactions	6
get_rd_credits	7
get_stamp_duty	8
get_tax_gap	9
get_tax_receipts	10
get_tobacco_duties	11
get_vat	12
list_tax_heads	13
tax_heads	14

Index	15
--------------	-----------

clear_cache	<i>Clear the local HMRC data cache</i>
-------------	--

Description

Deletes locally cached data files downloaded by the hmrc package. By default, all cached files are removed. Use `max_age_days` to remove only files older than a given number of days.

Usage

```
clear_cache(max_age_days = NULL)
```

Arguments

`max_age_days` Numeric or NULL. If NULL (default), all cached files are removed. If a number, only files last modified more than that many days ago are removed.

Value

Invisibly returns the number of files deleted.

Examples

```
# Remove all cached files
clear_cache()

# Remove files older than 30 days
clear_cache(max_age_days = 30)
```

get_corporation_tax *Download HMRC Corporation Tax receipts by type*

Description

Downloads and tidies the HMRC Corporation Tax Statistics annual publication, reporting receipts broken down by tax type for the most recent six financial years. Published annually in September.

Usage

```
get_corporation_tax(cache = TRUE)
```

Arguments

cache Logical. Use cached file if available (default TRUE).

Details

Note that some levies (e.g. Residential Property Developer Tax, Electricity Generators Levy) were introduced mid-series and will have NA values for earlier years.

Value

A data frame with columns:

tax_year Character. Financial year, e.g. "2023-24".

type Character. Tax type identifier.

description Character. Plain-English label.

receipts_gbp_m Numeric. Receipts in millions of pounds.

Source

<https://www.gov.uk/government/collections/analyses-of-corporation-tax-receipts-and-liabilities>

Examples

```
get_corporation_tax()
```

get_fuel_duties	<i>Download HMRC hydrocarbon oil (fuel duty) receipts</i>
-----------------	---

Description

Downloads and tidies the HMRC Hydrocarbon Oils Bulletin, which reports monthly fuel duty receipts. Data runs from January 1990 to the most recent published month, updated twice per year (January and July).

Usage

```
get_fuel_duties(fuel = NULL, start = NULL, end = NULL, cache = TRUE)
```

Arguments

fuel	Character vector or NULL (default = all). Valid values: "total", "petrol", "diesel", "other".
start	Character "YYYY-MM" or a Date object.
end	Character "YYYY-MM" or a Date object.
cache	Logical. Use cached file if available (default TRUE).

Value

A data frame with columns:

date Date. First day of the reference month.

fuel Character. Fuel category identifier.

description Character. Plain-English category label.

receipts_gbp_m Numeric. Duty receipts in millions of pounds.

Source

<https://www.gov.uk/government/statistics/hydrocarbon-oils-bulletin>

Examples

```
# Total fuel duty receipts since 2010
get_fuel_duties(fuel = "total", start = "2010-01")

# All categories
get_fuel_duties()
```

get_income_tax_stats *Download HMRC Income Tax liabilities by income range*

Description

Downloads and tidies HMRC Table 2.5, which reports the number of Income Tax payers and their liabilities grouped by total income range, for each available tax year. Numbers of taxpayers are in thousands; amounts are in millions of pounds unless otherwise noted. Published annually in May/June.

Usage

```
get_income_tax_stats(tax_year = NULL, cache = TRUE)
```

Arguments

tax_year Character vector or NULL (default = all years). Filter to specific tax years, e.g. "2023-24".

cache Logical. Use cached file if available (default TRUE).

Details

The earliest tax year with outturn data is based on the Survey of Personal Incomes; later years are projected estimates. Values suppressed for small sample sizes are returned as NA.

Value

A data frame with columns:

tax_year Character. Tax year, e.g. "2022-23".

income_range Character. Income range label.

income_lower_gbp Numeric. Lower limit of the income range in pounds. NA for the "All Ranges" row.

taxpayers_thousands Numeric. Number of Income Tax payers (thousands).

total_income_gbp_m Numeric. Total income (millions of pounds).

tax_liability_gbp_m Numeric. Total Income Tax liability (millions of pounds).

average_rate_pct Numeric. Average rate of Income Tax (percent).

average_tax_gbp Numeric. Average amount of Income Tax per taxpayer (pounds).

Source

<https://www.gov.uk/government/statistics/income-tax-liabilities-by-income-range>

Examples

```
get_income_tax_stats()

# Single tax year
get_income_tax_stats(tax_year = "2023-24")
```

```
get_property_transactions
```

Download monthly UK property transaction counts

Description

Downloads and tidies the HMRC Monthly Property Transactions bulletin, which counts residential and non-residential property transactions (SDLT returns, LBTT in Scotland, LTT in Wales) for England, Scotland, Wales, Northern Ireland, and the UK total. Data runs from April 2005 to the most recent completed month.

Usage

```
get_property_transactions(
  type = c("all", "residential", "non_residential"),
  nation = NULL,
  start = NULL,
  end = NULL,
  cache = TRUE
)
```

Arguments

<code>type</code>	Character. One of "all" (default), "residential", or "non_residential".
<code>nation</code>	Character vector or NULL (default = all nations). Valid values: "uk", "england", "scotland", "wales", "northern_ireland".
<code>start</code>	Character "YYYY-MM" or a Date object. Rows before this month are dropped.
<code>end</code>	Character "YYYY-MM" or a Date object. Rows after this month are dropped.
<code>cache</code>	Logical. If TRUE (default), the downloaded file is cached.

Value

A data frame with columns:

date Date. The first day of the reference month.

nation Character. One of "uk", "england", "scotland", "wales", or "northern_ireland".

type Character. "residential" or "non_residential".

transactions Numeric. Number of transactions.

Source

<https://www.gov.uk/government/statistics/monthly-property-transactions-completed-in-the-uk-with-va>

Examples

```
# All nations, all types
get_property_transactions()

# Residential only, England, since 2020
get_property_transactions(type = "residential", nation = "england",
                          start = "2020-01")
```

get_rd_credits	<i>Download HMRC R&D tax credit statistics</i>
----------------	--

Description

Downloads and tidies the HMRC Research and Development Tax Credits Statistics publication, covering the SME R&D Relief and Research and Development Expenditure Credit (RDEC) schemes. Annual data runs from 2000-01 to the most recent financial year, published annually in September.

Usage

```
get_rd_credits(scheme = NULL, measure = NULL, cache = TRUE)
```

Arguments

scheme	Character vector or NULL (default = all schemes). Valid values: "sme" (SME R&D Relief), "rdec" (Research and Development Expenditure Credit / large company scheme), "total" (combined).
measure	Character vector or NULL (default = all measures). Valid values: "claims" (number of claims), "amount_gbp_m" (cost in millions of pounds).
cache	Logical. Use cached file if available (default TRUE).

Details

Data before 2003-04 covers only the SME scheme (RDEC / large company scheme was introduced in 2002). Figures for the most recent two years are provisional and subject to revision as late claims are processed.

Value

A data frame with columns:

tax_year Character. Financial year, e.g. "2023-24".

scheme Character. Scheme identifier.

description Character. Plain-English scheme label.

measure Character. Either "claims" or "amount_gbp_m".

value Numeric. Number of claims or cost in millions of pounds.

Source

<https://www.gov.uk/government/statistics/corporate-tax-research-and-development-tax-credit>

Examples

```
# All R&D credit data
get_rd_credits()

# SME scheme claims only
get_rd_credits(scheme = "sme", measure = "claims")
```

get_stamp_duty	<i>Download HMRC stamp duty receipts</i>
----------------	--

Description

Downloads and tidies the HMRC Annual Stamp Tax Statistics, covering Stamp Duty Land Tax (SDLT), Stamp Duty Reserve Tax (SDRT) on shares, and other stamp duties. Annual data from 2003-04 to the most recent financial year, published each December.

Usage

```
get_stamp_duty(type = NULL, cache = TRUE)
```

Arguments

type	Character vector or NULL (default = all types). Valid values: "sdl_t_property" (SDLT on property excluding new leases), "sdl_t_leases" (SDLT on new leases), "sdl_t_total" (all SDLT), "sdr_t" (Stamp Duty Reserve Tax on shares), "stamp_duty" (Stamp Duty on documents), "total".
cache	Logical. Use cached file if available (default TRUE).

Value

A data frame with columns:

tax_year Character. Financial year, e.g. "2023-24".

type Character. Stamp duty type identifier.

description Character. Plain-English label.

receipts_gbp_m Numeric. Receipts in millions of pounds, rounded to nearest £5m.

Source

<https://www.gov.uk/government/statistics/uk-stamp-tax-statistics>

Examples

```
# All stamp duty types
get_stamp_duty()

# SDLT only
get_stamp_duty(type = "sdlt_total")
```

get_tax_gap

Download HMRC tax gap estimates

Description

Downloads and tidies the HMRC Measuring the Tax Gap publication, which estimates the difference between the tax theoretically owed and the tax actually collected, broken down by tax type and taxpayer group. Published annually in June, covering the most recent financial year.

Usage

```
get_tax_gap(tax = NULL, cache = TRUE)
```

Arguments

tax Character vector or NULL (default = all taxes). Filter by tax type, e.g. "Income Tax", "VAT", "Corporation Tax". Use `unique(get_tax_gap())$tax` to see all available values.

cache Logical. Use cached file if available (default TRUE).

Details

The tax gap publication is cross-sectional: each edition covers a single financial year. This function returns data for the most recent edition available on GOV.UK. Historical estimates back to 2005-06 are available in a separate HMRC publication (Measuring the Tax Gap: Time Series).

Value

A data frame with columns:

tax_year Character. Financial year of the estimate, e.g. "2023-24".

tax Character. Tax type.

taxpayer_type Character. Taxpayer group (e.g. "Individuals", "Small businesses").

component Character. Behaviour component (e.g. "Evasion", "Error", "Avoidance").

gap_pct Numeric. Tax gap as a percentage of the theoretical tax liability. NA where not disclosed.

gap_gbp_bn Numeric. Absolute tax gap in billions of pounds.

uncertainty Character. HMRC uncertainty rating for the estimate (e.g. "Low", "Medium", "High").

Source

<https://www.gov.uk/government/statistics/measuring-tax-gaps>

Examples

```
# Full tax gap breakdown
get_tax_gap()

# VAT gap only
get_tax_gap(tax = "VAT")
```

get_tax_receipts

Download HMRC tax receipts and National Insurance contributions

Description

Downloads and tidies the monthly HMRC Tax Receipts and National Insurance Contributions bulletin published on GOV.UK. The bulletin covers all major UK taxes and duties from April 2016 to the most recent month (monthly granularity), updated on approximately the 15th working day of each month.

Usage

```
get_tax_receipts(tax = NULL, start = NULL, end = NULL, cache = TRUE)
```

Arguments

tax	Character vector of tax head identifiers, or NULL (default) to return all available series. Use <code>list_tax_heads()</code> to see valid values and descriptions.
start	Character "YYYY-MM" or a Date object. If provided, rows before this month are dropped.

end	Character "YYYY-MM" or a Date object. If provided, rows after this month are dropped.
cache	Logical. If TRUE (default), the downloaded file is cached locally and reused on subsequent calls. Use <code>clear_cache()</code> to reset.

Value

A data frame with columns:

date Date. The first day of the reference month.

tax_head Character. Tax or duty identifier (see `list_tax_heads()`).

description Character. Plain-English series label.

receipts_gbp_m Numeric. Cash receipts in millions of pounds (GBP).

Source

<https://www.gov.uk/government/statistics/hmrc-tax-and-nics-receipts-for-the-uk>

Examples

```
# All tax heads
get_tax_receipts()

# Income Tax and VAT only
get_tax_receipts(tax = c("income_tax", "vat"))

# Since April 2020
get_tax_receipts(start = "2020-04")

# VAT in a specific window
get_tax_receipts(tax = "vat", start = "2019-01", end = "2024-12")
```

get_tobacco_duties *Download HMRC tobacco duty receipts*

Description

Downloads and tidies the HMRC Tobacco Bulletin, which reports monthly tobacco products duty receipts by product type. Data runs from January 1991 to the most recent published month, updated twice per year (February and August).

Usage

```
get_tobacco_duties(product = NULL, start = NULL, end = NULL, cache = TRUE)
```

Arguments

product	Character vector or NULL (default = all products). Valid values: "cigarettes", "cigars", "hand_rolling", "other", "total".
start	Character "YYYY-MM" or a Date object.
end	Character "YYYY-MM" or a Date object.
cache	Logical. Use cached file if available (default TRUE).

Value

A data frame with columns:

date Date. First day of the reference month.

product Character. Product type identifier.

description Character. Plain-English product label.

receipts_gbp_m Numeric. Duty receipts in millions of pounds.

Source

<https://www.gov.uk/government/statistics/tobacco-bulletin>

Examples

```
# All products since 2015
get_tobacco_duties(start = "2015-01")

# Cigarettes only
get_tobacco_duties(product = "cigarettes")
```

get_vat

Download HMRC VAT receipts

Description

Downloads and tidies the HMRC VAT Annual Statistics bulletin, which reports monthly VAT receipts broken down into payments, repayments, import VAT, and home VAT. Monthly data runs from April 1973; the bulletin is published annually (December).

Usage

```
get_vat(measure = NULL, start = NULL, end = NULL, cache = TRUE)
```

Arguments

measure	Character vector or NULL (default = all measures). Valid values: "total", "payments", "repayments", "import_vat", "home_vat".
start	Character "YYYY-MM" or a Date object.
end	Character "YYYY-MM" or a Date object.
cache	Logical. Use cached file if available (default TRUE).

Details

Note that early years (pre-1985) have suppressed payment and repayment splits; only the total is available for those periods. From January 2021, import VAT collected via postponed VAT accounting is recorded within payments and repayments rather than the import VAT column.

Value

A data frame with columns:

date Date. First day of the reference month.

measure Character. VAT measure identifier.

description Character. Plain-English measure label.

receipts_gbp_m Numeric. Value in millions of pounds. Repayments are negative (money flowing out from HMRC to businesses).

Source

<https://www.gov.uk/government/statistics/value-added-tax-vat-annual-statistics>

Examples

```
# Total VAT receipts since 2010
get_vat(measure = "total", start = "2010-01")

# Full breakdown
get_vat(start = "2020-01")
```

list_tax_heads

List available tax heads in the HMRC Tax Receipts bulletin

Description

Returns a data frame describing all tax and duty heads available in `get_tax_receipts()`. No network connection is required; the data is bundled with the package.

Usage

```
list_tax_heads()
```

Value

A data frame with columns:

tax_head Character. The identifier used in the tax argument of `get_tax_receipts()`.

description Character. Plain-English description of the series.

category Character. Broad grouping: "income", "expenditure", "consumption", "property", "environment", "nics", or "total".

available_from Character. Earliest year of monthly data (approximate).

Examples

```
list_tax_heads()
```

tax_heads	<i>Tax head lookup table</i>
-----------	------------------------------

Description

A data frame describing all tax and duty series available in `get_tax_receipts()`.

Usage

```
tax_heads
```

Format

A data frame with 29 rows and 4 columns:

tax_head Character. Identifier used in the tax argument of `get_tax_receipts()`.

description Character. Plain-English description.

category Character. Broad grouping: "income", "nics", "consumption", "property", "environment", "expenditure", "other", or "total".

available_from Character. Approximate start year of monthly data.

Source

Derived from the HMRC Tax Receipts and NICs bulletin. <https://www.gov.uk/government/statistics/hmrc-tax-and-nics-receipts-for-the-uk>

Index

* datasets

tax_heads, [14](#)

clear_cache, [2](#)

clear_cache(), [11](#)

get_corporation_tax, [3](#)

get_fuel_duties, [4](#)

get_income_tax_stats, [5](#)

get_property_transactions, [6](#)

get_rd_credits, [7](#)

get_stamp_duty, [8](#)

get_tax_gap, [9](#)

get_tax_receipts, [10](#)

get_tax_receipts(), [13](#), [14](#)

get_tobacco_duties, [11](#)

get_vat, [12](#)

list_tax_heads, [13](#)

list_tax_heads(), [10](#), [11](#)

tax_heads, [14](#)