# Package 'SouthKoreAPIs'

September 1, 2025

**Type** Package

**Title** Access South Korean Data via Public APIs and Curated Datasets

**Version** 0.1.0

**Maintainer** Renzo Caceres Rossi <arenzocaceresrossi@gmail.com>

**Description** Provides functions to access data from public RESTful APIs including 'Nager.Date', 'World Bank API', and 'REST Countries API', retrieving real-time or historical data related to South Korea, such as holidays, economic indicators, and international demographic and geopolitical indicators. Additionally, the package includes one of the largest curated collections of open datasets focused on South Korea, covering topics such as public health outbreaks, demographics, social surveys, elections, economic indicators, natural disasters, administrative divisions, air quality, climate data, energy consumption, cultural information, and financial markets. The package supports reproducible research and teaching by integrating reliable international APIs and structured datasets from public, academic, and government sources.
For more information on the APIs, see:
'Nager.Date' <https://date.nager.at/Api>,
'World Bank API' <https://datahelpdesk.worldbank.org/knowledgebase/articles/889392>,
and 'REST Countries API' <https://restcountries.com/>.

**License** MIT + file LICENSE

**Language** en

**URL** https://github.com/lightbluetitan/southkoreapis,
https://lightbluetitan.github.io/southkoreapis/

**BugReports** https://github.com/lightbluetitan/southkoreapis/issues

**Encoding** UTF-8

**LazyData** true

**Depends** R (>= 4.1.0)

**Imports** utils, httr, jsonlite, dplyr, scales, tibble

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

**RoxygenNote** 7.3.2

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Renzo Caceres Rossi [aut, cre] (ORCID:
<https://orcid.org/0009-0005-0744-854X>)

**Repository** CRAN

**Date/Publication** 2025-09-01 09:00:02 UTC

# Contents

## AutoOwnershipKorea_df *Korean Auto Ownership Data*

### Description

This dataset, AutoOwnershipKorea_df, is a data frame containing information on auto ownership in South Korea, along with related economic indicators. It includes data on gross national product, car prices, and oil prices over a series of years.

### Usage

```
data(AutoOwnershipKorea_df)
```

### Format

A data frame with 10 observations and 5 variables:

**Year** Year of observation (numeric)

**AO** Auto ownership level (numeric)

**GNP** Gross National Product (numeric)

**CP** Car price (numeric)

**OP** Oil price (numeric)

### Details

The dataset name has been kept as 'AutoOwnershipKorea_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame object. The original content has not been modified in any way.

### Source

Data taken from the **SenSrivastava** package version 2015.6.2

## demographicsKR_tbl_df *Korean Demographics (2000-2022)*

### Description

This dataset, demographicsKR_tbl_df, is a tibble containing demographic data of South Korea from 2000 to 2022. It includes values and rates for birth, death, natural growth, marriage, and divorce, organized by date and region. The dataset preserves the original structure from its source on Kaggle.

### Usage

```
data(demographicsKR_tbl_df)
```

## Format

A tibble with 4,860 observations and 12 variables:

**Date** Date of the record (character)

**Region** Region name in South Korea (character)

**Birth** Number of births (numeric)

**Birth_rate** Birth rate (per 1,000 people) (numeric)

**Death** Number of deaths (numeric)

**Death_rate** Death rate (per 1,000 people) (numeric)

**Divorce** Number of divorces (numeric)

**Divorce_rate** Divorce rate (per 1,000 people) (numeric)

**Marriage** Number of marriages (numeric)

**Marriage_rate** Marriage rate (per 1,000 people) (numeric)

**Natural_growth** Difference between births and deaths (numeric)

**Natural_growth_rate** Natural growth rate (per 1,000 people) (numeric)

## Details

The dataset name has been kept as 'demographicsKR_tbl_df' to maintain consistency with the naming conventions in the SouthKoreAPIs package. The suffix 'tbl_df' indicates that this is a tibble data frame. The original content has not been modified in any way.

## Source

Data obtained from Kaggle: https://www.kaggle.com/datasets/alexandrepetit881234/korean-demographics-200

---

# GasSales_Korea_tbl_df *Korea Natural Gas Sales with Temperature*

---

## Description

This dataset, GasSales_Korea_tbl_df, is a tibble containing monthly natural gas sales data with corresponding average temperatures for provinces of South Korea. It includes sales figures for each province, the national total, and temperature information, organized by year and month. The dataset preserves the original structure from its source on Kaggle.

## Usage

```
data(GasSales_Korea_tbl_df)
```

## Format

A tibble with 252 observations and 21 variables:

**Year** Year of observation (numeric)

**Month** Month of observation (numeric)

**Temperature** Average temperature in degrees Celsius (numeric)

**Gangwondo** Gas sales in Gangwon-do province (numeric)

**Seoul** Gas sales in Seoul (numeric)

**Gyeonggido** Gas sales in Gyeonggi-do province (numeric)

**Incheon** Gas sales in Incheon (numeric)

**Gyeongsangnamdo** Gas sales in Gyeongsangnam-do province (numeric)

**Gyeongsangbukdo** Gas sales in Gyeongsangbuk-do province (numeric)

**Gwangju** Gas sales in Gwangju (numeric)

**Daegu** Gas sales in Daegu (numeric)

**Daejeon** Gas sales in Daejeon (numeric)

**Busan** Gas sales in Busan (numeric)

**Sejong** Gas sales in Sejong (numeric)

**Ulsan** Gas sales in Ulsan (numeric)

**Jeollanamdo** Gas sales in Jeollanam-do province (numeric)

**Jeollabukdo** Gas sales in Jeollabuk-do province (numeric)

**Jeju** Gas sales in Jeju province (numeric)

**Chungcheongnamdo** Gas sales in Chungcheongnam-do province (numeric)

**Chungcheongbukdo** Gas sales in Chungcheongbuk-do province (numeric)

**Sum** Total gas sales in South Korea (numeric)

## Details

The dataset name has been kept as 'GasSales_Korea_tbl_df' to maintain consistency with the naming conventions in the SouthKoreAPIs package. The suffix 'tbl_df' indicates that this is a tibble data frame. The original content has not been modified in any way.

## Source

Data obtained from Kaggle: https://www.kaggle.com/datasets/zxtzxt30/korea-monthly-gas-sales-with-temper

`get_country_info_kr` *Get Country Information for South Korea*

## Description

Retrieves comprehensive country information for South Korea from the REST Countries API. This function fetches data including official and common names, geographical information, capital, area, population, and languages.

## Usage

```
get_country_info_kr()
```

## Details

This function makes a request to the REST Countries API v3.1 endpoint specifically for South Korea using full text search. It handles API errors gracefully and returns NULL if the request fails or no data is found.

## Value

A tibble with one row containing South Korea's country information:

**name_common** Common name of the country

**name_official** Official name of the country

**region** Geographic region

**subregion** Geographic subregion

**capital** Capital city(ies)

**area** Total area in square kilometers

**population** Total population

**languages** Languages spoken (comma-separated)

## Examples

```
# Get South Korea information
sk_info <- get_country_info_kr()
print(sk_info)
```

`get_southkorea_child_mortality` *Get South Korea's Under-5 Mortality Rate from World Bank*

## Description

Retrieves South Korea's under-5 mortality rate, measured as the number of deaths of children under five years of age per 1,000 live births, for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `SH.DYN.MORT`.

## Usage

```
get_southkorea_child_mortality()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Mortality rate, under-5 (per 1,000 live births)")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Mortality rate (per 1,000 live births)

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/SH.DYN.MORT

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
if (interactive()) {
  get_southkorea_child_mortality()
}
```

`get_southkorea_cpi` *Get South Korea's Consumer Price Index (2010 = 100) from World Bank*

## Description

Retrieves South Korea's Consumer Price Index (CPI), with 2010 as the base year (index = 100), for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `FP.CPI.TOTL`.

## Usage

```
get_southkorea_cpi()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Consumer price index (2010 = 100)")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Consumer Price Index (numeric, base year 2010 = 100)

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/FP.CPI.TOTL

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
if (interactive()) {
  get_southkorea_cpi()
}
```

`get_southkorea_energy_use` *Get South Korea's Energy Use (kg of oil equivalent per capita) from World Bank*

## Description

Retrieves South Korea's energy use per capita, measured in kilograms of oil equivalent, for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `EG.USE.PCAP.KG.OE`.

## Usage

```
get_southkorea_energy_use()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Energy use (kg of oil equivalent per capita)")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Energy use in kilograms of oil equivalent per capita

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
if (interactive()) {
  get_southkorea_energy_use()
}
```

`get_southkorea_gdp` *Get South Korea's GDP (current US$) from World Bank*

## Description

Retrieves South Korea's Gross Domestic Product (GDP) in current US dollars for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `NY.GDP.MKTP.CD`.

## Usage

```
get_southkorea_gdp()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "GDP (current US$)")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: GDP in current US dollars
- `value_label`: Formatted GDP with commas (e.g., "1,800,000,000,000")

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

## See Also

`GET`, `fromJSON`, `as_tibble`, `comma`

## Examples

```
if (interactive()) {
  get_southkorea_gdp()
}
```

`get_southkorea_holidays` *Get Official Public Holidays in South Korea for a Given Year*

## Description

Retrieves the list of official public holidays in South Korea for a specific year using the Nager.Date public holidays API. This function returns a tibble containing the date of the holiday, the name in the local language (Korean), and the English name. It is useful for academic, planning, and data analysis purposes. The information is retrieved directly from the Nager.Date API and reflects the current status of holidays for the requested year. The field names returned are consistent with the API structure.

## Usage

```
get_southkorea_holidays(year)
```

## Arguments

| | |
|---|---|
| `year` | An integer indicating the year (e.g., 2024 or 2025). |

## Value

A tibble with the following columns:

- `date`: Date of the holiday (class `Date`)
- `local_name`: Holiday name in the local language (Korean)
- `name`: Holiday name in English

## Source

Data obtained from the Nager.Date API: https://date.nager.at/

## Examples

```
get_southkorea_holidays(2024)
get_southkorea_holidays(2025)
```

`get_southkorea_hospital_beds` *Get South Korea's Hospital Beds (per 1,000 people) from World Bank*

## Description

Retrieves South Korea's number of hospital beds per 1,000 people for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `SH.MED.BEDS.ZS`.

## Usage

```
get_southkorea_hospital_beds()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Hospital beds (per 1,000 people)")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Number of hospital beds per 1,000 people

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/SH.MED.BEDS.ZS

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
if (interactive()) {
  get_southkorea_hospital_beds()
}
```

`get_southkorea_life_expectancy` *Get South Korea's Life Expectancy at Birth (Total, Years) from World Bank*

## Description

Retrieves South Korea's life expectancy at birth (total, years) for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `SP.DYN.LE00.IN`.

## Usage

```
get_southkorea_life_expectancy()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Life expectancy at birth, total (years)")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Life expectancy at birth in years

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/SP.DYN.LE00.IN

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
if (interactive()) {
  get_southkorea_life_expectancy()
}
```

`get_southkorea_literacy_rate` *Get South Korea's Adult Literacy Rate*

## Description

Retrieves South Korea's adult literacy rate ( for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `SE.ADT.LITR.ZS`.

## Usage

```
get_southkorea_literacy_rate()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Literacy rate as a percentage

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/SE.ADT.LITR.ZS

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
literacy_data <- get_southkorea_literacy_rate()
head(literacy_data)
```

`get_southkorea_population` *Get South Korea's Total Population from World Bank*

## Description

Retrieves South Korea's total population for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `SP.POP.TOTL`.

## Usage

```
get_southkorea_population()
```

## Details

The function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Population, total")
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Population as a numeric value
- `value_label`: Formatted population with commas (e.g., "51,000,000")

## Note

Requires internet connection. The data is retrieved in real time from the World Bank API.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/SP.POP.TOTL

## See Also

`GET`, `fromJSON`, `as_tibble`, `comma`

## Examples

```
if (interactive()) {
  get_southkorea_population()
}
```

`get_southkorea_unemployment` *Get South Korea's Unemployment Rate (Total) from World Bank*

## Description

Retrieves South Korea's total unemployment rate, measured as a percentage of the total labor force, for the years 2010 to 2022 using the World Bank Open Data API. The indicator used is `SL.UEM.TOTL.ZS`.

## Usage

```
get_southkorea_unemployment()
```

## Details

This function sends a GET request to the World Bank API. If the API request fails or returns an error status code, the function returns `NULL` with an informative message.

## Value

A tibble with the following columns:

- `indicator`: Indicator name (e.g., "Unemployment, total (
- `country`: Country name ("Korea, Rep.")
- `year`: Year of the data (integer)
- `value`: Unemployment rate as a numeric value (percentage)

## Note

Requires internet connection.

## Source

World Bank Open Data API: https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS

## See Also

`GET`, `fromJSON`, `as_tibble`

## Examples

```
if (interactive()) {
  get_southkorea_unemployment()
}
```

`HeptathlonSeoul1988_df`

*Olympic Heptathlon Results - Seoul 1988*

## Description

This dataset, HeptathlonSeoul1988_df, is a data frame containing the results of the Olympic heptathlon competition held in Seoul in 1988. It includes performance metrics for each of the seven events as well as the total score.

## Usage

```
data(HeptathlonSeoul1988_df)
```

## Format

A data frame with 25 observations and 8 variables:

**hurdles** Time in seconds for the 100m hurdles (numeric)

**highjump** Height in meters for the high jump (numeric)

**shot** Distance in meters for the shot put (numeric)

**run200m** Time in seconds for the 200m run (numeric)

**longjump** Distance in meters for the long jump (numeric)

**javelin** Distance in meters for the javelin throw (numeric)

**run800m** Time in seconds for the 800m run (numeric)

**score** Total heptathlon score (integer)

## Details

The dataset name has been kept as 'HeptathlonSeoul1988_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a base R data frame object. The original content has not been modified in any way.

## Source

Data taken from the **HSAUR3** package version 1.0-15

`KoreanBoneDensity_df` *Bone quality in South Koreans*

## Description

This dataset, KoreanBoneDensity_df, is a data frame containing bone mass density measurements of South Korean subjects at three body locations. It includes demographic information such as sex, age, height, and weight, along with bone mass density values for the lumbar spine, hip, and neck.

## Usage

```
data(KoreanBoneDensity_df)
```

## Format

A data frame with 969 observations and 7 variables:

**Sex** Sex of the subject (factor with 2 levels)

**Age** Age of the subject in years (integer)

**Height** Height of the subject in centimeters (numeric)

**Weight** Weight of the subject in kilograms (numeric)

**LumbarBMD** Bone mass density at the lumbar spine (numeric)

**HipBMD** Bone mass density at the hip (numeric)

**NeckBMD** Bone mass density at the neck (numeric)

## Details

The dataset name has been kept as 'KoreanBoneDensity_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame object. The original content has not been modified in any way.

## Source

Data taken from the **SRMData** package version 1.0.1

`KoreanElection2017_df` *2017 Korea Presidential Election Data*

## Description

This dataset, KoreanElection2017_df, is a data frame containing information from the 2017 presidential election in South Korea. It includes precinct- and city-level data along with demographic and socioeconomic indicators related to the voting population.

## Usage

```
data(KoreanElection2017_df)
```

## Format

A data frame with 1250 observations and 9 variables:

**PrecinctCode** Precinct code (integer)

**CityCode** City code (integer)

**CandidateName** Candidate identifier code (integer)

**AveAge** Average age of the voting population (numeric)

**AveYearEdu** Average years of education (numeric)

**AveHousePrice** Average house price (numeric)

**AveInsurance** Average insurance enrollment indicator or count (integer)

**VoteRate** Voter turnout rate (numeric)

**NumVote** Number of votes cast (integer)

## Details

The dataset name has been kept as 'KoreanElection2017_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame object. The original content has not been modified in any way.

## Source

Data taken from the **KPC** package version 0.1.2

`KoreanSocialSurvey_tbl_df` *A Sample of Korean General Social Survey Data, 2023*

## Description

This dataset, KoreanSocialSurvey_tbl_df, is a tibble containing a sample of data from the Korean General Social Survey (KGSS) conducted in 2023. It includes demographic, social, and attitudinal variables for respondents.

## Usage

```
data(KoreanSocialSurvey_tbl_df)
```

## Format

A tibble with 1123 observations and 13 variables:

**year** Survey year (numeric)

**respid** Respondent identifier (numeric)

**age** Age of the respondent (numeric)

**female** Gender indicator: 1 = female, 0 = male (numeric)

**employed** Employment status indicator (numeric)

**unived** University education indicator (numeric)

**netuse** Internet use indicator (numeric)

**ideo** Political ideology score (numeric)

**si_gbh** Regional code or classification (numeric)

**satisfin** Satisfaction with financial situation (numeric)

**fp_mord** Attitude toward moral issues (numeric)

**fpcat** Category for family planning or related topics (character)

**cntryaffq** Country affiliation or related attitude (character)

## Details

The dataset name has been kept as 'KoreanSocialSurvey_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble object. The original content has not been modified in any way.

## Source

Data taken from the **simqi** package version 0.2.0

## KOSPI200_list *Korea Stock Price Index 200 (KOSPI 200)*

### Description

This dataset, KOSPI200_list, is a list containing historical data for the Korea Stock Price Index 200 (KOSPI 200). It includes a vector of dates and the corresponding index values over time.

### Usage

```
data(KOSPI200_list)
```

### Format

A list with 2 components:

**date** A Date vector of length 896 representing the observation dates

**index** A numeric vector of length 896 representing the KOSPI 200 index values

### Details

The dataset name has been kept as 'KOSPI200_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'list' indicates that the dataset is stored as a list object. The original content has not been modified in any way.

### Source

Data taken from the **EMD** package version 1.5.9

## KPopIdols_tbl_df *All KPop Idols*

### Description

This dataset, KPopIdols_tbl_df, is a tibble containing a complete detailed database of all current KPop idols, both male and female. It includes each idol's stage name, full name, Korean name, Korean stage name, date of birth, group name, country, height, weight, birthplace, gender, and Instagram handle. The dataset preserves the original structure from its source on Kaggle.

### Usage

```
data(KPopIdols_tbl_df)
```

## Format

A tibble with 1,666 observations and 12 variables:

**Stage Name Stage Name** Stage name of the idol (character)

**Full Name Full Name** Full name of the idol (character)

**Korean Name Korean Name** Korean name of the idol (character)

**K. Stage Name K. Stage Name** Stage name written in Korean (character)

**Date of Birth Date of Birth** Date of birth (character)

**Group Group** Name of the group the idol belongs to (character)

**Country Country** Country of origin (character)

**Height Height** Height in centimeters (numeric)

**Weight Weight** Weight in kilograms (numeric)

**Birthplace Birthplace** Place of birth (character)

**Gender Gender** Gender of the idol (character)

**Instagram Instagram** Instagram handle or profile URL (character)

## Details

The dataset name has been kept as 'KPopIdols_tbl_df' to maintain consistency with the naming conventions in the SouthKoreAPIs package. The suffix 'tbl_df' indicates that this is a tibble data frame. The original content has not been modified in any way.

## Source

Data obtained from Kaggle: https://www.kaggle.com/datasets/onlyrohit/all-kpop-idols

---

# MERSKorea2015_list *Middle East respiratory syndrome in South Korea, 2015*

---

## Description

This dataset, MERSKorea2015_list, is a list containing two data frames with information collected during the first weeks of the Middle East respiratory syndrome (MERS-CoV) outbreak in South Korea in 2015. The data was initially gathered by the Epidemic Intelligence group at the European Centre for Disease Prevention and Control (ECDC).

## Usage

```
data(MERSKorea2015_list)
```

## Format

A list of 2 elements:

**linelist** A data frame with 162 observations and 15 variables:

- **id** Case identifier (character)
- **age** Age of the individual (integer)
- **age_class** Age class of the individual (character)
- **sex** Sex of the individual (factor with 2 levels)
- **place_infect** Place of infection (factor with 2 levels)
- **reporting_ctry** Reporting country (factor with 2 levels)
- **loc_hosp** Location or hospital (factor with 13 levels)
- **dt_onset** Date of symptom onset (Date)
- **dt_report** Date of case report (Date)
- **week_report** Week of report (factor with 5 levels)
- **dt_start_exp** Start date of exposure (Date)
- **dt_end_exp** End date of exposure (Date)
- **dt_diag** Date of diagnosis (Date)
- **outcome** Outcome of the case (factor with 2 levels)
- **dt_death** Date of death, if applicable (Date)

**contacts** A data frame with 98 observations and 4 variables:

- **from** ID of the source case (character)
- **to** ID of the contact case (character)
- **exposure** Type of exposure (factor with 5 levels)
- **diff_dt_onset** Difference in days between onset dates (integer)

## Details

The dataset name has been kept as 'MERSKorea2015_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'list' indicates that the object is a list containing multiple related data frames. The original content has not been modified in any way.

## Source

Data taken from the **outbreaks** package version 1.9.0

`migrationflows_tbl_df` *Annual Origin–Destination Migration Flows Between Korean Regions*

## Description

This dataset, migrationflows_tbl_df, is a tibble containing annual migration flows between South Korea's first-level administrative regions from 2012 to 2020. It includes geographic, economic, and demographic indicators for both origin and destination regions.

## Usage

```
data(migrationflows_tbl_df)
```

## Format

A tibble with 2,601 observations and 20 variables:

**orig** Origin region name (character)

**dest** Destination region name (character)

**year** Year of migration flow (integer)

**flow** Number of migrants moving from origin to destination (integer)

**dist_cent** Distance between region centroids (numeric)

**dist_min** Minimum distance between regions (numeric)

**dist_pw** Pairwise distance measure (numeric)

**contig** Contiguity indicator: TRUE if regions share a border, FALSE otherwise (logical)

**orig_pop** Population of the origin region (numeric)

**dest_pop** Population of the destination region (numeric)

**orig_area** Area of the origin region in square meters (numeric)

**dest_area** Area of the destination region in square meters (numeric)

**orig_gdp_pc** GDP per capita in the origin region (numeric)

**orig_ginc_pc** Gross income per capita in the origin region (numeric)

**orig_iinc_pc** Individual income per capita in the origin region (numeric)

**orig_pconsum_pc** Private consumption per capita in the origin region (numeric)

**dest_gdp_pc** GDP per capita in the destination region (numeric)

**dest_ginc_pc** Gross income per capita in the destination region (numeric)

**dest_iinc_pc** Individual income per capita in the destination region (numeric)

**dest_pconsum_pc** Private consumption per capita in the destination region (numeric)

## Details

The dataset name has been kept as 'migrationflows_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble object. The original content has not been modified in any way.

## Source

Data taken from the **migest** package version 2.0.5

---

`NFIColumnNames_df` *Korean and English Column Names*

---

## Description

This dataset, NFIColumnNames_df, is a data frame containing Korean and English translations of column names, along with their standardized column identifiers. It is intended to assist users in mapping between Korean-language variables and their English equivalents in related datasets.

## Usage

```
data(NFIColumnNames_df)
```

## Format

A data frame with 174 observations and 3 variables:

**Korean_Column_Name** Column name in Korean (character)

**English_Name** Column name in English (character)

**Column_Name** Standardized column identifier (character)

## Details

The dataset name has been kept as 'NFIColumnNames_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame object. The original content has not been modified in any way.

## Source

Data taken from the **knfi** package version 1.0.1.9

`RegionalKorea_df` *Korean Regional Data (2014–2016 Averages)*

## Description

This dataset, RegionalKorea_df, is a data frame containing average regional-level socioeconomic, demographic, health, and environmental indicators for South Korea over the period 2014–2016.

## Usage

```
data(RegionalKorea_df)
```

## Format

A data frame with 268 observations and 23 variables:

**id** Region identifier (integer)

**metro** Metropolitan area indicator or name (character)

**region** Region name (character)

**type** Region type classification (integer)

**grdp** Gross Regional Domestic Product (numeric)

**regpop** Regional population (numeric)

**popgrowth** Population growth rate (numeric)

**eq5d** EQ-5D health index (numeric)

**deaths** Number of deaths (numeric)

**drink** Alcohol consumption rate (numeric)

**hdrink** Heavy drinking rate (numeric)

**smoke** Smoking rate (numeric)

**aged** Proportion of elderly population (numeric)

**divorce** Divorce rate (numeric)

**medrate** Medical service utilization rate (numeric)

**gcomp** Gini coefficient or related inequality measure (numeric)

**vehipc** Vehicles per capita (numeric)

**accpv** Accidents per vehicle (numeric)

**dumppc** Illegal dumping incidents per capita (numeric)

**stratio** Student–teacher ratio (numeric)

**deathrate** Death rate (numeric)

**pctmale** Percentage of male population (numeric)

**accpc** Accidents per capita (numeric)

### Details

The dataset name has been kept as 'RegionalKorea_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame object. The original content has not been modified in any way.

### Source

Data taken from the **loedata** package version 1.0.1

---

## SeoulAdminAreas_sf *Administrative Areas of Seoul, South Korea*

---

### Description

This dataset, SeoulAdminAreas_sf, is an `sf` object containing polygon geometries for the 25 administrative areas of Seoul, Republic of Korea. It includes the area names, associated integer values, and polygon geometry data.

### Usage

```
data(SeoulAdminAreas_sf)
```

### Format

An `sf` object (tibble) with 25 observations and 3 variables:

**name** Name of the administrative area (character)

**value** Associated value or identifier (integer)

**geometry** Polygon geometry data (`sfc_POLYGON`)

### Details

The dataset name has been kept as 'SeoulAdminAreas_sf' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'sf' indicates that the dataset is a simple features object. The original content has not been modified in any way.

### Source

Data taken from the **valuemap** package version 2.0.4

## SeoulDistrictPop_df *Seoul's Population and Area Data for Districts (2012)*

### Description

This dataset, SeoulDistrictPop_df, is a data frame containing the 2012 population and area data for each of the districts in the city of Seoul, South Korea. It also includes information on the founding year of each district.

### Usage

```
data(SeoulDistrictPop_df)
```

### Format

A data frame with 25 observations and 5 variables:

**District** Name of the district (character)

**City** Name of the city (character)

**Pop.2012** Population in 2012 (integer)

**Area** Area of the district in square kilometers (numeric)

**Founded** Year the district was founded (character)

### Details

The dataset name has been kept as 'SeoulDistrictPop_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a base R data frame object. The original content has not been modified in any way.

### Source

Data taken from the **micromapST** package version 3.0.4

## SeoulH3Data_tbl_df *H3 Addresses within Seoul, South Korea*

### Description

This dataset, SeoulH3Data_tbl_df, is a tibble containing H3 resolution level 8 addresses within Seoul, Republic of Korea, along with associated numeric values. The H3 geospatial indexing system is used for representing hexagonal cells covering the area of Seoul.

### Usage

```
data(SeoulH3Data_tbl_df)
```

### Format

A tibble with 1,329 observations and 2 variables:

**name** H3 address at resolution level 8 (character)

**value** Associated numeric value (numeric)

### Details

The dataset name has been kept as 'SeoulH3Data_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble object. The original content has not been modified in any way.

### Source

Data taken from the **valuemap** package version 2.0.4

---

## SeoulMosquito_tbl_df *Mosquito Indicator in Seoul, Korea (2016-2019)*

---

### Description

This dataset, SeoulMosquito_tbl_df, is a tibble containing daily mosquito indicator data and weather measurements for Seoul, South Korea, from 2016 to 2019. The mosquito indicator represents the number of mosquitoes per specific area. The dataset also includes daily precipitation and temperature metrics (mean, minimum, and maximum). The dataset preserves the original structure from its source on Kaggle.

### Usage

```
data(SeoulMosquito_tbl_df)
```

### Format

A tibble with 1,342 observations and 6 variables:

**date** Observation date (Date)

**mosquito_Indicator** Number of mosquitoes per specific area (numeric)

**rain(mm)** Daily precipitation in millimeters (numeric)

**mean_T(°C)** Mean daily temperature in degrees Celsius (numeric)

**min_T(°C)** Minimum daily temperature in degrees Celsius (numeric)

**max_T(°C)** Maximum daily temperature in degrees Celsius (numeric)

### Details

The dataset name has been kept as 'SeoulMosquito_tbl_df' to maintain consistency with the naming conventions in the SouthKoreAPIs package. The suffix 'tbl_df' indicates that this is a tibble data frame. The original content has not been modified in any way.

### Source

Data obtained from Kaggle: https://www.kaggle.com/datasets/kukuroo3/mosquito-indicator-in-seoul-korea

---

## SolarRadiation_df *Solar Radiation Observations in South Korea*

---

### Description

This dataset, SolarRadiation_df, is a data frame containing hourly solar radiation measurements recorded at three locations in South Korea: Seoul, Daegu, and Busan. Observations cover the period from September 1, 2003, to September 29, 2003, and were obtained from the Korea Meteorological Administration.

### Usage

```
data(SolarRadiation_df)
```

### Format

A data frame with 696 observations and 4 variables:

**Date** Date and time of observation (POSIXct)

**Seoul** Solar radiation in Seoul (numeric)

**Daegu** Solar radiation in Daegu (numeric)

**Busan** Solar radiation in Busan (numeric)

### Details

The dataset name has been kept as 'SolarRadiation_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame object. The original content has not been modified in any way.

### Source

Data taken from the **EPT** package version 0.7.6

SouthKoreaBirths_tbl_df *Births in South Korea*

## Description

This dataset, SouthKoreaBirths_tbl_df, is a tibble containing births and mid-year population data for South Korea by age of mother, region, and calendar year from 2011 to 2023. It also includes regional data on GDP per capita (2023) and population density (2020).

## Usage

```
data(SouthKoreaBirths_tbl_df)
```

## Format

A tibble with 1,872 observations and 7 variables:

**age** Age group of the mother (character)

**region** Region name (factor with 16 levels)

**time** Calendar year (integer)

**births** Number of births (integer)

**popn** Mid-year population (integer)

**gdp_pc_2023** GDP per capita in 2023 (numeric)

**dens_2020** Population density in 2020 (character)

## Details

The dataset name has been kept as 'SouthKoreaBirths_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble object. The original content has not been modified in any way.

## Source

Data taken from the **bage** package version 0.9.4

SouthKoreaCovid19_tbl_df

*South Korea COVID-19 dataset*

## Description

This dataset, SouthKoreaCovid19_tbl_df, is a tibble containing COVID-19 data for South Korea from January 20th 2019 to March 20th 2020. It includes epidemiological, demographic, and geographic variables.

## Usage

```
data(SouthKoreaCovid19_tbl_df)
```

## Format

A tibble with 244 observations and 11 variables:

**n_covid1** Number of COVID-19 cases (numeric)

**Morbidity** Morbidity rate (numeric)

**high_sch_p** Proportion of population with high school education (numeric)

**Healthcare_access** Healthcare access index (numeric)

**diff_sd** Difference in standard deviation or related metric (numeric)

**Crowding** Crowding index (numeric)

**Migration** Migration rate (numeric)

**Health_behavior** Health behavior index (numeric)

**x** Longitude coordinate (numeric)

**y** Latitude coordinate (numeric)

**ln_total** Natural log of total cases or population (numeric)

## Details

The dataset name has been kept as 'SouthKoreaCovid19_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the SouthKoreAPIs package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a tibble object. The original content has not been modified in any way.

## Source

Data taken from the **gwzinbr** package version 0.1.0

## SouthKoreAPIs — *SouthKoreAPIs: Access South Korean Data via Public APIs and Curated Datasets*

### Description

This package provides functions to access data from public RESTful APIs including 'Nager.Date', 'World Bank API', and 'REST Countries API', retrieving real-time or historical data related to South Korea, such as holidays, economic indicators, and international demographic and geopolitical indicators. Additionally, the package includes one of the largest curated collections of open datasets focused on South Korea, covering topics such as public health outbreaks, demographics, social surveys, elections, economic indicators, natural disasters, administrative divisions, air quality, climate data, energy consumption, cultural information, and financial markets.

### Details

SouthKoreAPIs: Access South Korean Data via Public APIs and Curated Datasets

Access South Korean Data via Public APIs and Curated Datasets.

### Author(s)

**Maintainer**: Renzo Caceres Rossi <arenzocaceresrossi@gmail.com>

### See Also

Useful links:

- https://github.com/lightbluetitan/southkoreapis

## view_datasets_SouthKoreAPIs — *View Available Datasets in SouthKoreAPIs*

### Description

This function lists all datasets available in the 'SouthKoreAPIs' package. If the 'SouthKoreAPIs' package is not loaded, it stops and shows an error message. If no datasets are available, it returns a message and an empty vector.

### Usage

```
view_datasets_SouthKoreAPIs()
```

### Value

A character vector with the names of the available datasets. If no datasets are found, it returns an empty character vector.

## Examples

```
if (requireNamespace("SouthKoreAPIs", quietly = TRUE)) {
  library(SouthKoreAPIs)
  view_datasets_SouthKoreAPIs()
}
```

# Index